6-17-2019

Combinatorial biosynthesis and the basis for substrate promiscuity in class I diterpene synthases

Meirong Jia
Iowa State University

Sambit K. Mishra
Iowa State University

Samuel Tufts
Iowa State University

Robert L. Jernigan
Iowa State University, jernigan@iastate.edu

Reuben J. Peters
Iowa State University, rjpeters@iastate.edu

Combinatorial biosynthesis and the basis for substrate promiscuity in class I diterpene synthases

Abstract

Terpene synthases are capable of mediating complex reactions, but fundamentally simply catalyze lysis of allylic diphosphate esters with subsequent deprotonation. Even with the initially generated tertiary carbocation this offers a variety of product outcomes, and deprotonation further can be preceded by the addition of water. This is particularly evident with labdane-related diterpenes (LRDs) where such lysis follows bicyclization catalyzed by class II diterpene cyclases (DTCs) that generates preceding structural variation. Previous investigation revealed that two diterpene synthases (DTSs), one bacterial and the other plant-derived, exhibit extreme substrate promiscuity, but yet still typically produce *exo*-ene or tertiary alcohol LRD derivatives, respectively (i.e., demonstrating high catalytic specificity), enabling rational combinatorial biosynthesis. Here two DTSs that produce either *cis* or *trans endo*-ene LRD derivatives, also plant and bacterial (respectively), were examined for their potential analogous utility. Only the bacterial *trans-endo*-ene forming DTS was found to exhibit significant substrate promiscuity (with moderate catalytic specificity). This further led to investigation of the basis for substrate promiscuity, which was found to be more closely correlated with phylogenetic origin than reaction complexity. Specifically, bacterial DTSs exhibited significantly more substrate promiscuity than those from plants, presumably reflecting their distinct evolutionary context. In particular, plants typically have heavily elaborated LRD metabolism, in contrast to the rarity of such natural products in bacteria, and the lack of potential substrates presumably alleviates selective pressure against such promiscuity. Regardless of such speculation, this work provides novel biosynthetic access to almost 19 LRDs, demonstrating the power of the combinatorial approach taken here.

Disciplines

Biochemistry | Biophysics | Molecular Biology | Structural Biology

Comments

This is a manuscript of an article published as Jia, Meirong, Sambit K. Mishra, Samuel Tufts, Robert L. Jernigan, and Reuben J. Peters. "Combinatorial biosynthesis and the basis for substrate promiscuity in class I diterpene synthases." *Metabolic Engineering* (2019). doi: 10.1016/j.ymben.2019.06.008. Posted with permission.

Combinatorial biosynthesis and the basis for substrate promiscuity in class I diterpene synthases

Meirong Jia[1], Sambit K. Mishra[2], Samuel Tufts, Robert L. Jernigan, Reuben J. Peters*

Roy J. Carver Department of Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames, IA 50011, USA

*Corresponding author: e-mail address: rjpeters@iastate.edu

[1]Current address: Department of Plant Biology, University of California, Davis, Davis, CA 95616, USA

[2]Current address: Division of Biological Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Highlights

1. Extended studies of the extreme promiscuity exhibited by the bacterial *exo*-ene producing KgTS and plant derived tertiary alcohol producing SsSS with additional potential substrates.
2. Investigation of *endo*-ene yielding diterpene synthases identifies the highly promiscuous bacterial *trans*-*endo*-ene producing ScLS.
3. Further investigation of the basis for substrate promiscuity reveals that this is more correlated with phylogenetic origin than reaction mechanism, with bacterial diterpene synthases exhibiting significantly less selectivity than those from plants.
4. The substrate promiscuity found here enables combinatorial biosynthesis, providing novel access to almost 19 labdane-related diterpenes.

Abstract

Terpene synthases are capable of mediating complex reactions, but fundamentally simply catalyze lysis of allylic diphosphate esters with subsequent deprotonation. Even with the initially generated tertiary carbocation this offers a variety of product outcomes, and deprotonation further can be preceded by the addition of water. This is particularly evident with labdane-related diterpenes (LRDs) where such lysis follows bicyclization catalyzed by class II diterpene cyclases (DTCs) that generates preceding structural variation. Previous investigation revealed that two diterpene synthases (DTSs), one bacterial and the other plant-derived, exhibit extreme substrate promiscuity, but yet still typically produce *exo*-ene or tertiary alcohol LRD derivatives, respectively (i.e., demonstrating high catalytic specificity), enabling rational combinatorial biosynthesis. Here two DTSs that produce either *cis* or *trans endo*-ene LRD derivatives, also plant and bacterial (respectively), were examined for their potential analogous utility. Only the bacterial *trans-endo*-ene forming DTS was found to exhibit significant substrate promiscuity (with moderate catalytic specificity). This further led to investigation of the basis for substrate promiscuity, which was found to be more closely correlated with phylogenetic origin than reaction complexity. Specifically, bacterial DTSs exhibited significantly more substrate promiscuity than those from plants, presumably reflecting their distinct evolutionary context. In particular, plants typically have heavily elaborated LRD metabolism, in contrast to the rarity of such natural products in bacteria, and the lack of potential substrates presumably alleviates selective pressure against such promiscuity. Regardless of such speculation, this work provides novel biosynthetic access to almost 19 LRDs, demonstrating the power of the combinatorial approach taken here.

Abbreviations:

LRDs, labdane-related diterpenes;
DTCs, diterpene cyclases;
DTSs, diterpene synthases;
IDS, isoprenyl diphosphate synthases;
KgTS, terpentetriene synthase from *Kitasatospora griseola*;
SsSS, sclareol synthase from *Salvia sclarea*;
AbCAS, *cis*-abienol synthase from *Abies balsamea*;
SmMS, miltiradiene synthase from *Salvia miltiorrhiza*;
SaPS, diterpene synthase from marine bacterium *Salinispora arenicola*;
BjKS, *ent*-kaurene synthase from *Bradyrhizobium japonicum*;
ScLS, labda-8(17),12*E*,14-triene synthase from *Streptomyces cyslabdanicus* K04-0144;
EtKS, *ent*-kaurene synthase from *Erwinia tracheiphila*;
OsKS, *ent*-kaurene synthase from *Oryza sativa*;
AtKS, *ent*-kaurene synthase from *Arabidopsis thaliana*;
GGPP, (*E,E,E*)-geranylgeranyl diphosphate;
NNPP, (*Z,Z,Z*)-nerylneryl diphosphate;
GGPPS, GGPP synthase;
NNPPS, NNPP synthase;
FPP, (*E,E*)-farnesyl diphosphate;
GFPP, (*E,E,E,E*)-geranylfarnesyl diphosphate;
FPPS, FPP synthase;
GFPPS, GFPP synthase;
CPP, copalyl diphosphate;
AgAS, abietaenol synthase from *Abies grandis*;
SmCPS/KSL1, labda-7,13*E*-dien-15-ol synthase from *Selaginella moellendorffii*;
NgCLS, 8α-hydroxy-CPP synthase from *Nicotiana glutinosa*;
MvCPS1, peregrinol diphosphate synthase from *Marrubium vulgare*;
An2/ZmCPS2, *ent*-CPP synthase from *Zea mays*;
AtCPS, *ent*-CPP synthase from *Arabidopsis thaliana*;
OsCPS4, *syn*-CPP synthase from *Oryza sativa*;
Haur_2145, kolavenyl diphosphate synthase from *Herpetosiphon aurantiacus*;
KgTPS, terpentedienyl diphosphate synthase from *Kitasatospora griseola*;
MtHPS, tuberculosinyl/halimadienyl diphosphate synthase from *Mycobacterium tuberculosis*;
KPP, kolavenyl diphosphate;
syn-HPP, *syn*-halimadienyl diphosphate;
IPTG, isopropylthiogalactoside;
RT, retention time;
MS, mass spectra.

1. Introduction

While often associated with highly complex cyclization and rearrangement reactions, as suggested by their nomenclature terpene synthases (TPSs) do not necessarily catalyze such reactions. Essentially, their basic catalytic function is simply lysis of the allylic diphosphate ester, accomplished with assistance from a trio of divalent magnesium (Mg^{2+}) co-factors bound by conserved DDxxD and (N/D)Dxx(S/T)xxxE motifs, as well as conserved basic residues, followed by deprotonation (Christianson, 2008). Indeed, a significant portion of the complexity associated with TPS-mediated reactions can be attributed to the inherent reactivity of their isoprenoid substrates (Tantillo, 2017). The simplest TPS reaction is direct deprotonation of the initially generated tertiary carbocation to generate an olefin, although preceding addition of water can lead to formation of a tertiary alcohol as well. There are 3 potential olefin products, the *exo*-ene generated by deprotonation of the neighboring methyl, along with the *cis* and *trans* variants of the *endo*-ene generated by deprotonation of the neighboring methylene, which depends on the orientation of the allylic diphosphate isoprenyl unit relative to the rest of the precursor (Fig. 1).

Diterpenes, composed of four isoprenyl units, are generally derived from (*E,E,E*)-geranylgeranyl diphosphate (GGPP, **1**), although it has been shown that the *cisoid* analog (*Z,Z,Z*)-nerylneryl diphosphate (NNPP, **2**) also can serve as a precursor (Zi et al., 2014b). Such metabolism is particularly prevalent in plants, where GGPP is required to form photosynthetic pigments (i.e., the phytol side-chain of chlorophyll as well as carotenoids). Moreover, GGPP must be cyclized to produce the gibberellin A (GA) phytohormone required for normal growth and development in all vascular plants. The relevant DTC and DTS, which produce 8(17)-ene(*exo*) copalyl diphosphate (CPP) and kaurene, are termed CPP synthases (CPSs) and kaurene synthases (KSs), respectively. These have given rise to diversified families of DTCs and DTSs (these latter

sometimes referred to as KS-like or KSL) in many plant species, with the DTSs further serving as the ancestors of the plant TPS family more generally (Zi et al., 2014a). Accordingly, GA serves as the ancestral LRD and has given rise to extensive such metabolism in the plant kingdom. By contrast, LRD biosynthesis is only intermittently found, even rare, in microbes, such as fungi and bacteria (Zi et al., 2014a).

The LRDs are defined by the initiating bicyclization reaction mediated by DTCs (Peters, 2010). These catalyze protonation of the terminal carbon-carbon double bond (C=C), with *anti* addition of the two internal C=C leading to formation of the eponymous labda-13*E*-en-8-yl$^+$ diphosphate intermediate. This intermediate can be formed in four different stereochemical configurations, depending on the initial conformation of the substrate, which can be distinguished by the configuration of carbons 9 and 10 (C9 and C10; see Fig. 2 for numbering). These are traditionally designated as normal (9*S*,10*S*), *ent*- (9*R*,10*R*), *syn*- (9*R*,10*S*) or *ent-syn*- (9*S*,10*R*) – e.g., the gibberellins are derived from the *ent* stereoisomer. Note that the decalin bridgehead configuration (i.e., relative orientation of the C5-hydrogen and C10-methyl) is always *trans* in this intermediate. Immediate deprotonation leads to formation of the corresponding stereoisomer of CPP. However, this intermediate also can undergo 1,2-shifts of the hydride and methyl substituents of the tertiary and quaternary carbons (respectively) in the decalin bicycle, creating a series of tertiary carbocations. The first methyl shift creates the halimane backbone (halima-13*E*-en-10-yl$^+$ diphosphate), and the second generates the clerodane backbone, also referred to as kolavane (kolava-13*E*-en-4-yl$^+$ diphosphate), which is used here to distinguish the resulting product (KPP) from CPP (Fig. 2). Note that this final 1,2-shift can occur with either methyl substituent of C4, so the final decalin bridgehead (C5,10) configuration can be either *cis* or *trans*, adding further stereochemical variation. Each carbocation can be deprotonated at alternative

positions, such that each of these three basic backbones (labdane, halimane and clerodane) can be produced as 2-3 distinct olefins (i.e., C=C positioning), yielding isomers of CPP, HPP or KPP (respectively). In addition, each tertiary carbocation intermediate can undergo addition of water prior to deprotonation, yielding the corresponding hydroxylated derivative. Moreover, it has recently been demonstrated that DTCs can mediate ring rearrangement of the decalin bicycle as well (Xu et al., 2018). Accordingly, there are a wide variety of potential DTC products, although enzymes for only 19 are currently known (Tables 1 and S1). These DTC products serve as intermediates that undergo further biosynthetic elaboration, generally initiated by DTSs.

While DTS can catalyze elaborate reactions, as exemplified by that mediated by KSs, these require cyclization and, thus, rely on precise positioning of the internal C=C or hydroxyl group relative to the initially generated (tertiary) carbocation (Christianson, 2008). By contrast, the simplest (non-cyclizing) TPS reactions described above only require steric isolation of the allylic diphosphate isoprenyl unit. Accordingly, such DTSs might exhibit less stringent substrate binding – here termed substrate promiscuity, with fidelity in product outcome referred to as catalytic specificity, as previously defined (Hult and Berglund, 2007). Particularly if these DTSs exhibit significant substrate promiscuity yet are still catalytically specific, they can predictably build on the structural complexity mediated by the preceding DTCs, offering the possibility of rational combinatorial biosynthesis to generate substantial additional variety. This has been partially realized in previous work that identified two such DTSs that orthogonally catalyze two of the possible 'simple' product outcomes. In particular, the terpentetriene synthase from the bacterium *Kitasatospora griseola* (Dairi et al., 2001; Hamano et al., 2002), termed here KgTS, and sclareol synthase from the plant *Salvia sclarea* (Caniard et al., 2012; Schalk et al., 2012), termed here SsSS, which were found to react with a wide range of DTC products (i.e., all 12 known at that time),

generally yielding the *exo*-ene or tertiary alcohol LRD derivatives (i.e., 13(16)-ene or 13-ol), respectively (Jia et al., 2016). This work was enabled by a previously developed modular metabolic engineering system (Cyr et al., 2007), which enables facile co-expression in *Escherichia coli* of combinations of the relevant isoprenyl diphosphate synthase (IDS) and/or DTC and DTS, with additional engineering to increase metabolic flux to the upstream isoprenoid precursors (Morrone et al., 2010), such that sufficient quantities of the resulting product can be readily isolated for *de novo* structural characterization. Here this approach was used to further investigate the utility of additional DTSs for such rational combinatorial biosynthesis by examining potential general production of either *cis*- or *trans*- *endo*-ene LRD derivatives, as well as to investigate the basis for promiscuity via similar analysis of a variety of other DTSs, testing not only the acyclic 20-carbon (20C) precursors **1** and **2**, but also the *transoid* 15C precursor (*E,E*)-farnesyl diphosphate (FPP, **3**) and 25C precursor (*E,E,E,E*)-geranylfarnesyl diphosphate (GFPP, **4**), along with an expanded arsenal of 15 DTCs with distinct products (**5** – **19**; Table 1).

2. Results and Discussion

2.1. Further examination of KgTS and SsSS promiscuity

It was previously reported that KgTS will react with not only the 20C *transoid* isoprenyl diphosphate precursor **1**, but also the *transoid* 15C acyclic precursor FPP (**3**), at least *in vitro* (Hamano et al., 2002). In the *E. coli* metabolic engineering system endogenous phosphatases dephosphorylate isoprenyl diphosphate precursors, yielding the corresponding primary alcohol derivative (designated here by prime notation of the corresponding compound number – i.e., **1'** – **19'**), which are extractable and observable by the gas-chromatograph with mass spectral (GC-MS) detection analytical method utilized here. To further examine the promiscuity of KgTS, as well as

SsSS, in the context of this system it seemed worth investigating their ability to react with *transoid* acyclic precursors that differ in length by a single isoprenyl unit – i.e., both the shorter **3** and longer **4**. This was carried out via co-expression with IDSs that produce either **3** or **4** (i.e., FPP synthase, FPPS, or GFPP synthase, GFPPS). As the native (plant) genes contain N-terminal plastidial targeting peptides that are removed after import *in planta*, the recombinant genes used here for GFPPS and SsSS are truncated to remove the corresponding sequence and, thus, encode pseudo-mature enzymes. Note that analogous pseudo-mature constructs are used for the plant-derived IDSs that produce either **1** or **2** (i.e., GGPP synthase, GGPPS, or NNPP synthase, NNPPS), as well as all plant-derived DTSs and DTCs.

In this metabolic engineering system both KgTS and SsSS were found to react with the shorter **3**, albeit somewhat inefficiently (i.e., at least relative to the *E. coli* phosphatases, as less than half of the observed diterpenoids result from DTS activity, with the remainder representing the primary alcohol derivative produced by the phosphatases), but only KgTS reacts with **4** and does so reasonably efficiently. Although it was previously reported that only SsSS reacted with the *cisoid* C20 acyclic precursor **2** (Jia et al., 2016), the greater promiscuity observed for KgTS with **4** prompted reexamination of its reactivity with the *cisoid* **2** – i.e., via co-expression with NNPPS (Zi et al., 2014b). Indeed, extending the data collection time to cover earlier eluting compounds revealed that KgTS does react with **2**, actually significantly more efficiently than does SsSS (Figs 3 and S1).

As previously reported from *in vitro* assays (Hamano et al., 2002), with **3** KgTS seems to produce a mixture of the three C=C isomers of farnesene (**20** – **22**), while SsSS more specifically produces the tertiary alcohol (*E*)-nerolidol (**23**), as verified by comparison via GC-MS of retention time (RT) and mass spectra (MS) to an authentic standard (Fig. S2). With **4** KgTS produces a

mixture of three sesterterpenes assumed to be the C=C isomers of geranylfarnesene (**24** – **26**), although the predominant **25**, based on relative RT and MS, appears to be the *trans*-*endo*-ene, commonly referred to as the (*E*)-α isomer (Fig. S2). With **2** KgTS yields two products, the major of which, upon isolation and structural analysis by NMR (Figs. S3-S5 and Table S2), was found to be the expected *exo*-ene containing derivative of **2**, β-nerylmyrcene (**27**). However, while another diterpene was observed as a minor product, surprisingly, upon isolation and structural analysis by NMR (Figs S6-S8 and Table S3), this was found to be β-springene (**28**). It is assumed that this is derived from production of (*E,E,Z*)-geranylneryl diphosphate from the endogenous FPP (**3**, from the *E. coli* host) by the NNPPS (with the putative dephosphorylated derivative of this indicated by P in Fig. 3). Regardless, the KgTS activity found here provides novel biosynthetic access to **27** and **25** (albeit alongside small amounts of C=C isomer co-products in each case).

As previously reported (Jia et al., 2016), although KgTS and SsSS react more efficiently with **1** than the endogenous phosphatases from *E. coli*, they cannot compete with DTCs. This enables facile examination of their ability to react with DTC produced bicyclic isoprenyl diphosphate precursors (i.e., by simple co-expression of the relevant DTC, as well as GGPPS). Such an approach was taken here to examine the specificity of KgTS and SsSS with three additional DTC products. These are an *ent*- version of 7-*endo*-CPP (*ent*-7-*endo*-CPP, **9**), a novel example of the decalin ring with bridgehead C=C, *syn*-halima-5(10),13*E*-dienyl diphosphate (**18**), and, of particular interest, a 5-6 bicycle resulting from ring rearrangement (of the halima-13*E*-en-5-yl$^+$ diphosphate intermediate) such that the product no longer has the prototypical decalin core, mutildienyl diphosphate (**19**). Both KgTS and SsSS react more efficiently than the endogenous *E. coli* phosphatases with **9** and **18**, but only KgTS does so with **19**, indicating that the bacterial KgTS exhibits more substrate promiscuity than the plant-derived SsSS (Figs 3 and S1).

Perhaps not surprisingly, the products of KgTS and SsSS with **9** appear to be *enantiomers* of their products with the normal stereoisomer (i.e., 7-*endo*-CPP, **8**), and, on the basis of their identical RT and MS in (non-chiral) GC-MS analysis, were assigned as the expected *exo*-ene derivative *ent*-labda-7,13(16),14-triene (**29**) and tertiary alcohol derivative *ent*-labda-7,14-dien-13-ol (**30**), respectively (Fig. S9). The KgTS and SsSS products with **18** were unknown and it was necessary to isolate these for *de novo* structural analysis by NMR, which determined that these are the expected *exo*-ene derivative (Figs. S10-S12 and Table S4), *syn*-halima-5(10),13(16),14-triene (**31**), and tertiary alcohol derivative (Figs. S13-S15 and Table S5), *syn*-halima-5(10),14-dien-13-ol (**32**), respectively. Although SsSS produces only **32**, KgTS produces both, with actually slightly more **32** than **31**. The KgTS product with **19** also was unknown, but upon isolation and *de novo* structural analysis by NMR was found to be the expected *exo*-ene derivative (Figs. S16-S18 and Table S6), termed here mutil-4(18),13(16),14-triene (**33**).

While **32** has been previously isolated from a plant extract (Nagashima et al., 2001), the relevant DTS is unknown. In addition, searches of the SciFinder database indicate that **29** – **31** and **33** do not appear to have been previously reported. Accordingly, the KgTS and SsSS activities reported here not only generally confirm their utility for rational combinatorial biosynthesis, but also provide novel biosynthetic access to **29** – **33** (albeit **31** is only produced in a mixture with **32**).

2.2. Examining the promiscuity of endo-ene producing DTSs

As described above, there are four potential product outcomes with the simplest TPS reaction. Accordingly, in addition to the *exo*-ene and tertiary alcohol produced by KgTS and SsSS (respectively), it should also be possible generate *endo*-ene derivatives, in either the *cis* or *trans* configuration. Indeed, there is a DTS known to catalyze each of these outcomes. In particular, the *cis*-abienol synthase from the plant *Abies balsamea* (Zerbe et al., 2012), termed here AbCAS,

which produces the *cis*-*endo*-ene derivative of its native substrate 8α-hydroxy-CPP (**10**), and a labdatriene synthase from the bacterium *Streptomyces cyslabdanicus* K04-0144 (Yamada et al., 2016), termed here ScLS, which produces the *trans*-*endo*-ene derivative of its native substrate CPP (**5**). Note that AbCAS is bifunctional, with both DTC and DTS activity, but a variant with a D405A substitution that negates DTC activity – i.e., this AbCAS construct only exhibits DTS activity – has been reported (Zerbe et al., 2012), and was used here.

Both AbCAS and ScLS were tested with all the potential substrates already examined with KgTS and SsSS (i.e., **1** – **19**). Strikingly, the plant-derived AbCAS was found to be quite specific, reacting efficiently only with its native substrate **10**, with the only other accepted precursor being the structurally closely related **5**, and even this reacts relatively poorly (Fig. 4 and Table S7). By contrast, the bacterial ScLS is highly promiscuous, reacting with all the potential substrates except the longer (25C) **4** (Fig. 5 and Table S7). While ScLS reacts relatively poorly with the acyclic substrates (i.e., **1** – **3**) and the 5-6 bicycle **19**, it efficiently out-competes the *E. coli* phosphatases with all the decalin containing substrates (i.e., **5** – **18**).

The observed MS for the products of AbCAS and ScLS with their native substrates are consistent with the previously reported activity – i.e., the production of *cis*-abienol (labda-12*Z*,14-dien-8α-ol, **34**) from **10** by AbCAS (Zerbe et al., 2012), and the production of labda-8(17),12*E*,14-triene (**35**) from **5** by ScLS (Yamada et al., 2016). Similarly, as previously reported (Yamada et al., 2016), ScLS efficiently reacts with **8** to specifically produce the *trans*-*endo*-ene derivative labda-7,12*E*,14-triene (**36**). The mass spectra for **34** – **36** are shown in the Supporting Information (Fig. S19).

In a number of cases the observed products seemed likely to correspond to the known products of KgTS, SsSS or other known DTSs, or be *enantiomers* of these, which was investigated

by GC-MS based comparison of RT and MS. For example, with its acyclic substrates ScLS was found to specifically produce the expected *trans*-*endo*-ene derivative from **1**, (*E*)-α-springene (**37**), as identified from the mix produced by KgTS (Nakano et al., 2010), but surprisingly specifically produce the same *exo*-ene derivative from **2** as KgTS (i.e., **27**), and a similar mixture of the C=C isomers of farnesene (**20** – **22**) as KgTS, as well as the SsSS product (*E*)-nerolidol (**23**), from **3**. Perhaps more interestingly, AbCAS yields two products from **5**, with the less abundant one found to be the same *exo*-ene derivative sclarene (**38**) produced by KgTS with **5**, while the major product was the expected *cis*-*endo*-ene derivative, (*Z*)-biformene (labda-8(17),12*Z*,14-triene, **39**), as verified by comparison to the *enantiomer* known to be produced from **6** by KgTS (Fig. S20). Similarly, ScLS yields two products from **6**, the *enantiomer* of its native substrate **5**, with the major product being the *exo*-ene derivative *ent*-sclarene (**40**), as identified by comparison to the *enantiomer* produced by KgTS with **5** (Fig. S21), while the minor product is the known *ent*-kaurene (**41**). Interestingly, KgTS does not produce **40** from **6**, but rather the *cis*-*endo*-ene derivative instead, such that ScLS and KgTS both exhibit unexpected yet orthologous activity with this substrate. With **7** ScLS also yields two products, which were found to be the same *exo*-ene derivative griseolaene (*syn*-labda-8(17),13(16),14-triene, **42**) produced by KgTS, and the same tertiary alcohol derivative vitexifolin A (*syn*-labda-8(17),14-dien-13-ol, **43**) produced by SsSS (Fig. S22). ScLS further yields two products from **9**, and the major product was found to be the same *exo*-ene derivative **29** produced by KgTS, while the minor product was the expected *trans*-*endo*-ene derivative, *ent*-labda-7,12*E*,14-triene (**44**), as identified by comparison to the *enantiomer* produced by ScLS with **8** (Fig. S22). With **11** ScLS efficiently and selectively yields a single product, identified as the heterocyclic derivative *ent*-13-*epi*-manoyl oxide (**45**)(Fig. S22), as also produced by KSs from this hydroxylated derivative of their native substrate (Mafu et al., 2015).

With **13** ScLS selectively produces the same tertiary alcohol derivative cleroda-3,14-dien-8-ol (**46**) as SsSS (Fig. S22). Finally, with **19** ScLS produces small amounts of the same *exo*-ene derivative **33** as KgTS, along with trace amounts of an unidentified diterpene (Fig. S22).

In the case of the remaining substrates, at least one of the observed ScLS products could not be identified by comparison to readily available diterpenes, requiring isolation and *de novo* structural analysis by NMR. For example, with **10** ScLS yields three products. By comparison to previously identified DTS products (Mafu et al., 2015), the two more abundant products were identified (Fig. S23) as manoyl oxide (**47**) and its C13 *epimer*, 13-*epi*-manoyl oxide (**48**), while the third required *de novo* structural analysis (Figs. S24-S26 and Table S8), and was identified as the expected *trans*-*endo*-ene derivative *trans*-abienol (labda-12*E*,14-dien-8α-ol, **49**). With **12** ScLS yields two products and, upon *de novo* structural analysis (Figs. S27-S32 and Tables S9-S10), the major product was identified as *syn*-labda-9,13*S*-epoxy-14-ene (**50**), while the minor product was identified as the C13 *epimer syn*-labda-9,13*R*-epoxy-14-ene (**51**). With **14** ScLS yields a single product that, upon *de novo* structural analysis (Figs. S33-S35 and Table S11), was identified as the expected *trans*-*endo*-ene derivative *ent*-cleroda-3,12*E*,14-triene (**52**). With **15** ScLS yields two products, the predominant of which, upon *de novo* structural analysis (Figs. S36-S38 and Table S12), was identified as the expected *trans*-*endo*-ene derivative *syn*-cleroda-3,12*E*,14-triene (**53**), while the minor product was identified as the same *exo*-ene derivative *syn*-cleroda-3,13(16),14-triene (**54**) produced by KgTS (Fig. S23). With **16** ScLS yields a single product that, upon *de novo* structural analysis (Figs. S39-S41 and Table S13), was identified as the expected *trans*-*endo*-ene derivative halima-5,12*E*,14-triene (**55**). With **17** ScLS yields two products, the major of which was identified as the same *exo*-ene derivative *syn*-halima-5,13(16),14-triene (**56**) produced by KgTS (Fig. S23), while the minor product, upon *de novo*

structural analysis (Figs. S42-S44 and Table S14), was identified as the expected *trans*-*endo*-ene derivative *syn*-halima-5,12*E*,14-triene (**57**). With **18** ScLS yields three products, the major of which, upon *de novo* structural analysis (Figs. S45-S47 and Table S15), was identified as the expected *trans*-*endo*-ene derivative *syn*-halima-5(10),12*E*,14-triene (**58**), while the minor products were found to be the same *exo*-ene derivative **31** produced by KgTS and tertiary alcohol derivative **32** produced by SsSS.

While **39** has been previously isolated (Noma et al., 1982), the relevant DTS is unknown. Moreover, although production of **50** and **51** was recently reported (Johnson et al., 2018), that for **40**, **44**, **52**, **53**, **55**, **57** and **58** does not appear to have been previously reported. Accordingly, the AbCAS and, particularly, ScLS activities reported here provides novel access to **39**, **40**, **44**, **52**, **53**, **55**, **57** and **58** (albeit **39** is only produced inefficiently, and **39**, **40**, **44**, **53**, **57** and **58** are produced non-selectively).

It is evident that ScLS exhibits extreme substrate promiscuity but at least moderately specific catalytic activity, enabling (semi)-rational combinatorial biosynthesis as shown above. However, despite their catalysis of analogously simple TPS reactions, the substrate promiscuity of ScLS contrasts with the much greater specificity exhibited by AbCAS. Together with the greater promiscuity observed with KgTS versus SsSS this suggests the hypothesis that that bacterial DTSs might generally prove to be more promiscuous than those from plants.

2.3. Examining the phylogenetic basis for DTS promiscuity

To examine the hypothesis that substrate promiscuity depends more on phylogenetic origin than reaction mechanism, a variety of additional DTSs were chosen for investigation of their substrate specificity. Notably, this has already been examined to some extent with a number of plant DTSs. Although much of that work focused on selectivity for the various stereoisomers of

(*exo*-)CPP (i.e., **5** – **7**), more extensive studies also have been reported (Andersen-Ranberg et al., 2016; Zerbe et al., 2013). Of particular relevance here, it has been shown that the DTS activity of the enzymes most closely related to AbCAS (i.e., those from other gymnosperms) exhibit (*exo*-)CPP stereospecificity and will only react with their native substrate **5**, but not **6** or **7**. This includes not only the more closely related bifunctional abietadiene synthases (Peters et al., 2000), which catalyze more complex cyclization and rearrangement reactions, but also pimaradiene synthases that catalyze more straightforward simple cyclization reactions (Hall et al., 2013). Similarly, all the investigated KSs from plant gibberellin biosynthesis, which catalyze a highly complex (bi)cyclization with subsequent ring rearrangement reaction and yet are ancestral to all plant DTSs (Zi et al., 2014a), also exhibit CPP stereospecificity, reacting only with their native substrate, the *enantiomer* **6** (Cui et al., 2015; Heskes et al., 2018; Irmisch et al., 2015; Jackson et al., 2014; Kumar et al., 2016; Pelot et al., 2018; Shimane et al., 2014; Wu et al., 2012; Xu et al., 2007; Zerbe et al., 2014; Zerbe et al., 2013; Zhou et al., 2012). It has been reported that at least some plant pimaradiene synthases will react with two stereoisomers of CPP (Morrone et al., 2011; Pelot et al., 2018; Zhou et al., 2012), and a number of DTSs that catalyze cyclization reactions have been shown to readily react with the hydroxylated variant of their native substrate to carry out heterocyclization, forming various isomers of manoyl oxide (Andersen-Ranberg et al., 2016; Brückner et al., 2014; Ignea et al., 2015; Mafu et al., 2015; Pateraki et al., 2014). This latter group of DTSs includes the miltiradiene synthase from *Salvia miltiorrhiza* (Gao et al., 2009), termed here SmMS, which is closely related to SsSS (>60% amino acid, aa, sequence identity), but carries out both cyclization and rearrangement of the initially formed pimarane backbone (i.e., produces an abietane), and was chosen for broader analysis of substrate specificity. In addition, two plant KSs required for gibberellin biosynthesis, that from *Arabidopsis thaliana* (Yamaguchi et al., 1998),

termed here AtKS, and rice (Sakamoto et al., 2004), *Oryza sativa* so termed here OsKS, as representative of dicots and monocots respectively, also were selected to reflect their ancestral role in evolution of the plant DTS family. While functionally analogous AtKS and OsKS share only ~42% aa sequence identity. There are relatively few bacterial DTSs and these have almost invariably just been coupled to the products of the DTC from the relevant biosynthetic operon, although a report on somewhat broader studies with KgTS *in vitro* prompted its use for combinatorial biosynthesis (Nakano et al., 2010). Hence, to examine a 'simple' cyclase, the pimaradiene synthase from the bacterium *Salinispora arenicola* (Xu et al., 2014), termed here SaPS, was selected, along with two bacterial KSs, both also involved in gibberellin biosynthesis, one from *Bradyrhizobium japonicum* (Morrone et al., 2009), termed here BjKS, and the other from *Erwinia tracheiphila* (Nagel and Peters, 2017), termed here EtKS, to examine the effect of increased reaction complexity and match the phylogenetic divergence of the pair of plant KSs (albeit these are more closely related, sharing ~51% aa sequence identity).

The substrate specificity of these six DTSs was investigated with all potential substrates **1** – **19** (Figs. 6-7; note that the product profiles for BjKS and EtKS are essentially identical, so only those for BjKS are shown here). Strikingly, it was immediately evident that the bacterial DTSs exhibit much greater promiscuity than those from plants, particularly when including all those investigated here, as, with the exception of SsSS, the plant derived DTSs react with only a couple of substrates, while those from bacteria react with almost all and do so reasonably efficiently with at least half of the decalin bicycle containing substrates (Tables 2 and S7). Indeed, in contrast to the plant KSs, which only efficiently react with their native substrate **6** and the hydroxylated variant **11**, the bacterial KSs, if anything, are actually more promiscuous than SaPS, which catalyzes a simpler reaction. Nevertheless, the similarly limited substrate specificity exhibited by

the plant-derived AtKS and OsKS, as well as promiscuity exhibited by the bacterial BjKS and EtKS, while distinct, is consistent with the analogous physiological function of each pair in the relevant biological kingdom. Intriguingly, while production of *ent*-kaurene requires complex (bi)cyclization and ring rearrangement, the promiscuity of the bacterial KSs indicates that this reaction need not be tightly chaperoned, which is consistent with the production of *ent*-kaurene by ScLS as well. Indeed, the ring rearrangement has been predicted to be concerted (albeit asynchronously) with secondary (tetra)cyclization (Hong and Tantillo, 2010). Nevertheless, this further emphasizes both the inherent reactivity of these isoprenyl reactants (Tantillo, 2017), and importance of catalytic base positioning to terminate the carbocationic (cascade) reactions initiated by TPSs (Pemberton and Christianson, 2016), as the lack of deprotonation of earlier intermediates presumably at least partially underlies the production of this complex diterpene by these promiscuous bacterial DTSs.

The observed mass spectra for the products of these DTSs with their native substrates are consistent with the previously reported native activity – i.e., the production of *ent*-kaurene (**41**) from **6** by all four KSs (Morrone et al., 2009; Nagel and Peters, 2017; Xu et al., 2007; Yamaguchi et al., 1998), production of miltiradiene (abieta-8,12-diene, **59**) from **5** by SmMS (Gao et al., 2009), and production of isopimara-8,15-diene (**60**) from **5** by SaPS (Xu et al., 2014). In addition, as previously reported (Mafu et al., 2015), SmMS reacts with **10**, the hydroxylated variant of its native substrate, and predominantly produces manoyl oxide (**47**), along with small amounts of 13-*epi*-manoyl oxide (**48**), while AtKS and OsKS react with **11**, the hydroxylated variant of their native substrate, and specifically produce *ent*-13-*epi*-manoyl oxide (**45**). The mass spectra for **59** – **60** can be found in the Supporting Information (Fig. S48).

In a number of cases at least some of the observed products seemed likely to correspond to the known products of KgTS or SsSS, or be *enantiomers* of these, which was investigated by GC-MS based comparison of retention time and mass spectra. For example, AtKS reacts (albeit somewhat inefficiently) with **9**, the 7-*endo* isomer of its native substrate, and the product was found to be the same *ent*-labda-7,13(16),14-triene (**29**) as identified here from KgTS. The bacterial DTS examined here (i.e., SaPS, BjKS and EtKS) also reacted with many substrates, generally producing similarly simple derivatives. For example, with **1** SaPS selectively produces the *exo*-ene derivative β-springene (**28**), while BjKS and EtKS produce both this and the tertiary alcohol derivative geranyllinalool (**61**). With **3** BjKS and EtKS selectively produce the tertiary alcohol derivative **23**. With **5** BjKS and EtKS yield two products, one of which was identified as the tertiary alcohol derivative manool (**62**), and the other product was identified as pimara-7,15-diene (**63**) by comparison to its *enantiomer*, which is a known DTS product (Zhou et al., 2012) (Fig. S49). With **6** SaPS yields two products, with the major one identified as the tertiary alcohol derivative *ent*-manool (**64**), while the minor product was identified as *ent*-pimara-8(14),15-diene (**65**), a known DTS product (Xu et al., 2007). With **8** SaPS, BjKS and EtKS all selectively produce the tertiary alcohol derivative labda-7,14-dien-13-ol (**66**). Analogous results were obtained with **9**, as all three bacterial DTSs again selectively produce the tertiary alcohol derivative *ent*-labda-7,14-dien-13-ol (**30**). With **12** SaPS selectively produces the tertiary alcohol derivative viteagnusin D (*syn*-labda-14-en-9,13-diol, **67**), while BjKS and EtKS also produce small amount of **67**, they largely produce the *exo*-ene derivative *syn*-labda-13(16),14-dien-9-ol (**68**). With **14** BjKS and EtKS yield two products, one is the tertiary alcohol derivative *ent*-cleroda-3,14-dien-13-ol (**69**), and the other product is the *exo*-ene derivative *ent*-cleroda-3,13(16),14-triene (**70**). With **17** SaPS, BjKS and EtKS all produce the tertiary alcohol derivative *syn*-halima-5,14-dien-

13-ol (**71**). Analogous results were obtained with **18**, as all three bacterial DTSs again selectively produce the same tertiary alcohol derivative **32** as identified here with SsSS. The mass spectra for **64** – **71** can be found in the Supporting Information (Fig. S50).

In most other cases all the products have undergone cyclization. For example, with **7** BjKS and EtKS yield two products, with the major one identified as *syn*-pimara-7,15-diene (**72**), and the minor product as *syn*-stemodene (**73**), both of which are known DTS products (Morrone et al., 2006; Wilderman et al., 2004). With **10**, while BjKS and EtKS produce an *epimeric* mixture of manoyl oxide, **47** and **48**, SaPS selectively produces **47**, much as described for other DTSs (Mafu et al., 2015). Conversely, with the *enantiomer* **11**, SaPS produces an *epimeric* mixture of *ent*-manoyl oxide (**74**), with slightly more *ent*-13-*epi*-manoyl oxide (**45**) than **74**, while BjKS and EtKS selectively produce **45**, much like the plant KSs (Mafu et al., 2015). The mass spectra for **72** – **74** can be found in the Supporting Information (Fig. S51). Finally, with **7** SmMS yields a single product that required *de novo* structural analysis and, thus (Figs. S52-S54 and Table S16), was identified as *syn*-abieta-9(11),12-diene (**75**), while SaPS yields two products, the minor of which was **73**, while the major product also required *de novo* structural analysis and, hence (Figs. S55-S57 and Table S17), was identified as *syn*-pimara-9(11),15-diene (**76**).

While **75** has been previously isolated (Sakurai et al., 1999), the relevant DTS is unknown. Moreover, **76** does not appear to have been previously reported. Accordingly, the DTS activities reported here provide novel biosynthetic access to **63**, **75**, and **76** (albeit **63** and **76** are produced non-selectively). More generally, while the bacterial DTS catalyzing more complex reactions exhibit high substrate promiscuity, this is coupled to weak catalytic specificity.

2.4 Examining the structural basis for DTS promiscuity

With the striking exception of SsSS, the plant DTSs examined here exhibit strong substrate selectivity. Much as previously observed with other plant DTSs (Andersen-Ranberg et al., 2016; Brückner et al., 2014; Ignea et al., 2015; Mafu et al., 2015; Morrone et al., 2011; Pateraki et al., 2014; Pelot et al., 2018; Zhou et al., 2012), these only react with structurally closely related isomers. For example, while they all seem to react with the hydroxylated variant of their native (*exo*-)CPP substrate (or the reverse in the case of AbCAS), only AtKS reacts with the corresponding stereoisomer of 7-*endo*-CPP (and relatively poorly at that), while here only SmMS reacts with an alternative stereoisomer of (*exo*-)CPP (although, as noted above, analogous relaxed but not fully promiscuous CPP stereoselectivity has been observed with a few plant pimaradiene synthases). By contrast, the bacterial DTSs uniformly exhibited substrate promiscuity, differing only in degree, regardless of reaction complexity (Table 2). Indeed, the BjKS and EtKS that catalyze the most complex reaction exhibit greater promiscuity than does SaPS, despite its simpler reaction mechanism. Nevertheless, a degree of selectivity is observed with the three bacterial DTSs that catalyze more complex reactions, which provides some insight into their substrate binding constraints. For example, the low catalytic efficiency observed with the 20C **1** and, for the bacterial KSs the 15C **3**, versus the lack of reactivity with the C25 **4**, indicates the importance of compactness and rigidity imparted by DTC bicyclization of **1**, while the lack of reactivity of the *cisoid* **2** indicates that the *trans* configuration of the C=C allylic to the diphosphate ester also is important. More specifically, the lack of reactivity with the 5-6 bicycle **19** highlights the role of the 6-6 fused decalin core. Notably, while the DTC products examined here all contain *trans*-decalin bicycles, based on the previous reports of SsSS promiscuity (Andersen-Ranberg et al., 2016; Jia et al., 2016), it has been shown to react with the products of several other newly identified DTCs (Johnson et al., 2018; Pelot et al., 2018), including *syn*-*cis*-*endo*-KPP, which indicates that

the decalin bridgehead configuration is not critical for substrate recognition by promiscuous DTSs. On the other hand, despite their promiscuity with all available (stereo)isomers of CPP (**5** – **9**), and their hydroxylated variants (**10** – **12**), the three bacterial DTSs that catalyze more complex reactions also do not react with normal or *syn*-KPP (**13** and **15**), nor normal HPP (**16**), although they do react with the C=C isomers of *syn*-HPP (**17** and **18**), with varying efficiency, and the bacterial KSs with *ent*-KPP (**14**). Accordingly, it appears that the consecutive methyl migrations away from the labdane configuration found in CPP (i.e., to generate HPP and then KPP) progressively decrease catalytic efficiency, which is coupled to some stereospecificity (at least for the bacterial KSs), in certain cases leading to complete loss of reactivity.

Fortuitously, crystal structures for BjKS have been reported (Liu et al., 2014), which enabled more detailed computational investigation of the basis for its relaxed substrate selectivity observed here. Although the structure of a substrate bound form was reported, **6** was soaked into preexisting crystals that were grown under conditions with high concentrations of tartrate, which chelates metals, so the resulting structure does not contain the requisite Mg^{2+} co-factors. Perhaps not surprisingly, this structure further did not appear to have a fully closed active site, not least as it also is missing the loop between the J and K helices (J-K loop) that normally folds over the cavity. In order to model a catalytically competent form, the BjKS with **6** structure was used as a starting point, with the well-defined substrate and Mg^{2+} bound structure for bornyl diphosphate synthase used as the template (Whittington et al., 2002). This allowed generation of a completely enclosed active site into which the necessary three Mg^{2+} could be docked (Fig. S58), enabling further docking of three potential substrates with distinct reactivity. In particular, the native (fully reactive) substrate *ent*-CPP isomer **6**, partially reactive *syn*-HPP isomer **17** and the unreactive *syn*-KPP isomer **15**. The hydrocarbon backbones of each of these appears to be bound in quite different

configurations (Fig. S59), which then affects diphosphate positioning and interactions with the Mg^{2+} co-factors and other polar contacts. Of particular interest is the number of contacts formed by a highly conserved arginine (Arg204), which is part of an extension of the second characteristic TPS motif – i.e., RLx(N/D)Dxx(S/T/G)xxx(E/D) – and is important for lysis of the diphosphate ester (Liu et al., 2014), as the number of contacts is sequentially decreased between these (potential) substrates in concert with their reactivity (Fig. 8). This then provides a rationale for the somewhat selective reactivity observed here with BjKS. Similarly, modeling with the functionally analogous but phylogenetically disparate EtKS (51% sequence identity) suggests that its analogous reactivity can be rationalized in the same fashion – e.g., their active site volumes are almost exactly identical, 1720 $Å^3$ versus 1714 $Å^3$, respectively (Fig. S60).

3. Conclusions

Here the hypothesis that DTS catalyzing simple lysis and immediate deprotonation reactions in LRD biosynthesis would prove to exhibit substrate promiscuity with catalytic specificity and, hence, enable rational combinatorial biosynthesis, was investigated. In particular, to supplement the earlier finding of such utility for the bacterial *exo*-ene derivative producing KgTS and plant derived SsSS that produces tertiary alcohol derivatives, the plant derived AbCAS that naturally produces a *cis*-*endo*-ene derivative and bacterial ScLS that yields *trans*-*endo*-ene derivatives were targeted here. The bacterial ScLS was found to exhibit extreme substrate promiscuity with at least moderate catalytic specificity, approximating that observed with KgTS and SsSS, including an extended palette of potential substrates, together providing novel biosynthetic access to 15 LRDs (all the DTS products can be found in Tables 3 and S18, with the corresponding chemical structures found in Figs. 9 and S61). However, the plant derived AbCAS exhibited quite strict substrate specificity. In conjunction with previous reports of plant DTS

substrate specificity, this suggested the alternative hypothesis that substrate promiscuity might be more tightly correlated with phylogenetic origin than reaction mechanism. Indeed, investigation of additional DTSs, including SmMS (which is closely related to SsSS), strongly supports this alternative hypothesis, with even the investigated bacterial KSs that catalyze a particularly complex (bi)cyclization and ring rearrangement reaction exhibiting a high degree of substrate and catalytic promiscuity, contrasting with the strict specificity found for the plant KSs. Analysis of substrate reactivity relationships, alongside docking studies with the structurally defined bacterial BjKS, provided some insight into the basis for the promiscuity found here, specifically suggesting that this is based on accommodation of potential substrates in reactive configurations, and further highlights the intrinsic reactivity of these isoprenoid substrates. Regardless of such mechanistic speculation, the results reported here suggest that bacterial DTSs will generally prove to be significantly more promiscuous than those from plants. This may be a function of the universal presence of extensive LRD metabolism, derived from the requisite biosynthesis of gibberellin hormones, in plants, which contrasts with the rarity of such metabolism in bacteria. As previously suggested (Tawfik, 2014), the absence of alternative substrates presumably alleviates selective pressure against promiscuity. Whatever the underlying rationale, it will be of interest to more thoroughly investigate the full array of potential substrates with not only the now known extremely promiscuous SsSS, KgTS and ScLS, but also bacterial DTSs more generally, to mine these for novel biosynthetic capacity.

4. Methods and Materials

4.1. General

Unless otherwise noted, chemicals were purchased from Fisher Scientific and molecular biology reagents, including synthetic genes, from Invitrogen. The *trans*-nerolidol standard was

purchased from Sigma-Aldrich. All constructs were verified by full sequencing of the inserted gene.

4.2. Recombinant constructs

The modularity of the metabolic system utilized here is based on the use of DEST cassettes that enable facile recombination via the Gateway (Invitrogen) cloning system (Cyr et al., 2007), particularly as inserted into the Duet (Novagen) series of vectors (Morrone et al., 2010). Accordingly, all DTCs and DTSs were first cloned into the pENTR/SD/D-TOPO vector, excluding the N-terminal plastidial targeting peptide sequence from the plant-derived genes. The DTSs were then typically recombined into a pDEST15 expression vector, although SsSS and SmMS were similarly inserted into pDEST14, while EtKS was directly inserted into the commercial expression vector Champion™ pET100/D-TOPO for convenience. DTCs were generally inserted into the DEST cassette found in the previously described pGG-DEST vector (Cyr et al., 2007). To screen DTSs activity against linear precursors, for **1** the DTS expression vectors were co-transformed with a previously described pGG vector (Cyr et al., 2007), for **3** with a previously described pMevT-MBIS vector (Martin et al., 2003), for **4** with a pGF vector constructed by sub-cloning a previously described GFPPS, specifically AtIDS9 (Nagel et al., 2015), into the NcoI/EcoRI restriction sites of pACYC-Duet1, while for **2** the DTSs were recombined from pENTR into a pNN-DEST vector for single plasmid co-expression as previously described for SsSS and KgTS (Jia et al., 2016). To increase metabolic flux towards terpenoids, several key genes from the endogenous isoprenoid precursor pathway were over-expressed using a previously reported pIRS plasmid (Morrone et al., 2010).

4.3. Metabolic Engineering

All metabolic engineering was carried out using the *E. coli* OverExpress C41 strain (Lucigen), and included pIRS as well as the relevant expression constructs (i.e., for the production of **1, 2, 3** or **4,** and/or a DTC, along with a DTS). For initial activity screening purpose, recombinant cultures were grown in 50 mL TB medium (pH = 7.0), with appropriate antibiotics, in 250 mL Erlenmeyer flasks. These cultures were first grown at 37 °C to mid-log phase (OD_{600} ~0.7), then the temperature dropped to 16 °C for 0.5 h before induction with 1 mM isopropylthiogalactoside (IPTG) and supplementation with 40 mM pyruvate and 1 mM $MgCl_2$. The induced cultures were grown for an additional 72 h before extraction with an equal volume of hexanes, with the organic phase then separated, and concentrated under N_2 when necessary for analysis.

4.4. Diterpene product analysis by GC-MS chromatography

GC-MS analyses were carried out using a Varian 3900 GC with a Saturn 2100T ion trap mass spectrometer in electron ionization (70 eV) mode, with an Agilent HP-5MS column (Agilent, 19091S-433), run with a 1.2 mL/min helium flow rate. Samples (1 uL) were injected in splitless mode by an 8400 autosampler with the injection port at 250 °C. The following temperature program was used: the oven temperature initially started at 50 °C, which was maintained for 3 min, and then increased at a rate of 15 °C/min to 300 °C, where it was held for another 3 min. Mass data was recorded by mass-to-charge ratio (*m/z*) values in a range from 90 to 650, starting from 13 min after sample injection until the end of the run. For detection of 15-carbon products, the helium flow rate was reduced to 1.0 mL/min, and the temperature program was modified to have a smaller increasing rate of 8 °C/min to 250 °C, and mass data collected from 6 to 30 min, with the range extended down to 40 (i.e., *m/z* from 40 to 650), while all other parameters were kept the same.

4.5. Diterpene production and purification

To obtain sufficient amount of new enzymatic products for NMR analysis, the bacterial cultures described above were simply scaled up to 1 L in 2.8 L Fernbach flasks. All other procedures were identical except that the extraction was repeated twice to ensure full yield. The separated organic phase was pooled and dried by rotary evaporation under vacuum, and the residue was re-suspended in 5 mL hexane for subsequent fractionation via flash chromatography over a 4 g-silica column (Grace) using a Grace Reveleris flash chromatography system with UV detection and automated injector and fraction collector, run at 15 mL/min. Briefly, the column was pre-equilibrated with hexanes and the sample injected, followed by 100% hexane (0-4 min), 0-100% acetone (4-5 min), 100% acetone (5-8 min), with peak-based fraction collection (15 mL maximum per tube). Generally, non-oxygen containing products eluted in the 100% hexane fraction; otherwise, the products were found in the 100% acetone fractions. Fractions containing the diterpene product, as identified by GC-MS analysis, were dried under N_2, re-suspended in 2 mL methanol, and filtered through 0.2 um cellulose filter (Thermo Scientific). These fractions were further separated using an Agilent 1200 series HPLC instrument equipped with a diode array UV detector and automated injector and fraction collector, over a semi-preparative C-8 column (ZORBAX Eclipse XDB-C8, 25 × 0.94 cm) run at 4 mL/min. The column was pre-equilibrated with acetonitrile/water (1:1 for oxygenated products, 4:1 for olefins), the sample injected, followed by washing (0 – 2 min) with same acetonitrile/water mix (i.e., depending on the targeted compound), then the percentage of acetonitrile increased to 100% (2–10 min), and final elution with 100% acetonitrile (10–30 min). Peaks were collected and analyzed by GC-MS and, if necessary, the diterpene further purified by another round of HPLC separation over an analytical C-8 column (Kromasil® C8, 150 × 4.6 mm) run at 0.5 mL/min, and using the same elution program

described above. Fractions containing pure compounds were dried under N_2, and the compounds then dissolved in 0.5 mL deuterated $CDCl_3$ (Aldrich) for NMR analysis.

4.6. Chemical structure identification by NMR analysis

NMR spectra were acquired on a Bruker AVIII-800 spectrometer equipped with a 5-mm HCN cryogenic probe, set at 25 °C, using TopSpin 3.2 software. Chemical shifts were calculated by reference to those known for $CDCl_3$ (^{13}C 77.23 ppm, ^{1}H 7.24 ppm) signals offset from TMS. All spectra were acquired using standard programs from the TopSpin 3.2 software, with collection of 1D ^{1}H-NMR, and 2D double-quantum filtered correlation spectroscopy (DQF-COSY), heteronuclear single-quantum coherence (HSQC), heteronuclear multiple-bond correlation (HMBC), HMQC-COSY and NOESY (800 MHz), as well as 1D ^{13}C-NMR (201 MHz) spectra. Observed HMBC correlations were used to propose a partial structure, while COSY correlations between protonated carbons were used to complete the structure, which was further verified by HSQC correlations. Observed correlations from NOESY spectrum were used to assign the relative stereochemistry of chiral carbons and also the configuration of double bonds, where applicable. Absolute stereochemistry was assigned based on the known configuration of the upstream DTC product. The resulting structures were searched against the SciFinder database (https://scifinder.cas.org) to determine precedent.

4.7. Computational biology

The reported crystal structure of BjKS bound with substrate *ent*-CPP (PDB: 4XLY) is missing the J-K loop (residues 212-220) and also lacking the three Mg^{2+} ions required for catalysis. Therefore, the fully closed structure of BjKS was first modeled by addition of these two missing elements using the substrate analog and Mg^{2+} containing structure of bornyl diphosphate synthase (PDB: 1N20) as the template. From the structural alignment of 4XLY and 1N20 using Pymol, it

was found that residues 501-509 from 1N20 aligned quite well with the missing loop in 4XLY. The PDB coordinates for this region (residues 501-509, 1N20) were obtained from the template PDB file, and Modeller (Sali et al., 1997; Šali et al., 2014) used to build back in the missing functional loop of BjKS. This initial model was optimized using the Chiron energy minimization server (Ramachandran et al., 2011), and the resulting structure used for subsequent docking studies. This was initiated by docking of the three catalytically requisite Mg^{2+} ions using AutoDock and AutoDock Vina (Morris et al., 2009; Trott and Olson, 2010), specifically so that Mg^{2+}_a and Mg^{2+}_c interact with the DDXXD motif, particularly the first and last aspartates (i.e., D75 and D79), while Mg^{2+}_b interacts with the (N/D)Dxx(S/T/G)xxx(E/D) motif, particularly the first and last residues (i.e., N207 and D215), as almost universally found in active forms of TPSs (Christianson, 2008). The docking ligands were prepared in chem3D 15.1 with energy minimization using the MM2 molecular mechanics force field with all parameters set to their default values. The ligands then were docked individually into the modeled BjKS+Mg structure using AutoDock Vina. Twenty poses for each ligand were generated, with the lowest energy pose chosen for comparative purposes. Polar contacts to the diphosphate from both BjKS and the Mg^{2+} ions were identified using Pymol. Both the number and the type of polar contacts found in each representative pose were used to estimate the strength of interaction between the ligand and the enzyme. A structure for the (distantly) related EtKS was generated based on the original BjKS structure (PDB: 4XLY) using Modeller, and the closed structure constructed, again based on bornyl diphosphate synthase (PDB: 1N20), with final energy minimization via Chiron, as described above. This EtKS structure had a root mean-squared deviation (RMSD) of 1.2 Å with the modeled BjKS structure when considering all the residues. The volume of the active site pocket was then calculated using Chimera (Pettersen et al., 2004), based on key active site residues

in BjKS (F72, L68, I36, Y168, I166, A167, L71, Y136, D79, D76, D75, D208, R204, N207) and EtKS (F72, M68, I36, Y168, I166, G167, L71, Y136, D79, D76, D75, D208, R204, N207).

Acknowledgment

The authors thank Dr. Raimund Nagel for construction of the pGF vector. This work was supported by a grant from the NIH (GM109773 to R.J.P.), and grant from the NSF (DBI-1661391 to R.L.J.), as well as funds from the Roy J. Carver Trust awarded to the authors department.

References

Andersen-Ranberg, J., Kongstad, K. T., Nielsen, M. T., Jensen, N. B., Pateraki, I., Bach, S. S., Hamberger, B., Zerbe, P., Staerk, D., Bohlmann, J., Moller, B. L., Hamberger, B., 2016. Expanding the Landscape of Diterpene Structural Diversity through Stereochemically Controlled Combinatorial Biosynthesis. Angew Chem Int Ed Engl. 55**,** 2142-6.

Brückner, K., Bozic, D., Manzano, D., Papaefthimiou, D., Pateraki, I., Scheler, U., Ferrer, A., de Vos, R. C., Kanellis, A. K., Tissier, A., 2014. Characterization of two genes for the biosynthesis of abietane-type diterpenes in rosemary (Rosmarinus officinalis) glandular trichomes. Phytochemistry. 101**,** 52-64.

Burke, C., Croteau, R., 2002. Geranyl diphosphate synthase from Abies grandis: cDNA isolation, functional expression, and characterization. Archives of biochemistry and biophysics. 405**,** 130-6.

Caniard, A., Zerbe, P., Legrand, S., Cohade, A., Valot, N., Magnard, J. L., Bohlmann, J., Legendre, L., 2012. Discovery and functional characterization of two diterpene synthases for sclareol biosynthesis in *Salvia sclarea* (L.) and their relevance for perfume manufacture. BMC Plant Biol. 12**,** 119.

Christianson, D. W., 2008. Unearthing the roots of the terpenome. Curr Opin Chem Biol. 12**,** 141-50.

Criswell, J., Potter, K., Shephard, F., Beale, M. H., Peters, R. J., 2012. A single residue change leads to a hydroxylated product from the class II diterpene cyclization catalyzed by abietadiene synthase. Org Lett. 14**,** 5828-31.

Cui, G., Duan, L., Jin, B., Qian, J., Xue, Z., Shen, G., Snyder, J. H., Song, J., Chen, S., Huang, L., Peters, R. J., Qi, X., 2015. Functional divergence of diterpene synthases in the medicinal plant *Salvia miltiorrhiza*. Plant Physiol. 169**,** 1607-1618.

Cyr, A., Wilderman, P. R., Determan, M., Peters, R. J., 2007. A Modular Approach for Facile Biosynthesis of Labdane-Related Diterpenes. J. Am. Chem. Soc. 129**,** 6684-6685.

Dairi, T., Hamano, Y., Kuzuyama, Y., Itoh, N., Furihata, K., Seto, H., 2001. Eubacterial diterpene

cyclase genes essential for production of the isoprenoid antibiotic terpentecin. J. Bact. 183**,** 6085-6092.

Gao, W., Hillwig, M. L., Huang, L., Cui, G., Wang, X., Kong, J., Yang, B., Peters, R. J., 2009. A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. Org. Lett. 11**,** 5170-5173.

Hall, D. E., Zerbe, P., Jancsik, S., Quesada, A. L., Dullat, H., Madilao, L. L., Yuen, M., Bohlmann, J., 2013. Evolution of conifer diterpene synthases: diterpene resin acid biosynthesis in lodgepole pine and jack pine involves monofunctional and bifunctional diterpene synthases. Plant Physiol. 161**,** 600-16.

Hamano, Y., Kuzuyama, Y., Itoh, N., Furihata, K., Seto, H., Dairi, T., 2002. Functional analysis of eubacterial diterpene cyclases responsible for biosynthesis of a diterpene antibiotic, terpentecin. J. Biol. Chem. 277**,** 37098-37104.

Harris, L. J., Saparno, A., Johnston, A., Prisic, S., Xu, M., Allard, S., Kathiresan, A., Ouellet, T., Peters, R. J., 2005. The maize An2 gene is induced by Fusarium attack and encodes an ent-copalyl diphosphate synthase. Plant Mol Biol. 59**,** 881-894.

Heskes, A. M., Sundram, T. C. M., Boughton, B. A., Jensen, N. B., Hansen, N. L., Crocoll, C., Cozzi, F., Rasmussen, S., Hamberger, B., Hamberger, B., Staerk, D., Moller, B. L., Pateraki, I., 2018. Biosynthesis of bioactive diterpenoids in the medicinal plant Vitex agnus-castus. Plant J. 93**,** 943-958.

Hong, Y. J., Tantillo, D. J., 2010. Formation of beyerene, kaurene, trachylobane, and atiserene diterpenes by rearrangements that avoid secondary carbocations. J Am Chem Soc. 132**,** 5375-86.

Hult, K., Berglund, P., 2007. Enzyme promiscuity: mechanism and applications. Trends in biotechnology. 25**,** 231-8.

Ignea, C., Ioannou, E., Georgantea, P., Loupassaki, S., Trikka, F. A., Kanellis, A. K., Makris, A. M., Roussis, V., Kampranis, S. C., 2015. Reconstructing the chemical diversity of labdane-type diterpene biosynthesis in yeast. Metabolic engineering. 28**,** 91-103.

Ikeda, H., Shin-Ya, K., Nagamitsu, T., Tomoda, H., 2016. Biosynthesis of mercapturic acid derivative of the labdane-type diterpene, cyslabdan that potentiates imipenem activity against methicillin-resistant Staphylococcus aureus: cyslabdan is generated by mycothiol-mediated xenobiotic detoxification. J Ind Microbiol Biotechnol. 43**,** 325-42.

Irmisch, S., Muller, A. T., Schmidt, L., Gunther, J., Gershenzon, J., Kollner, T. G., 2015. One amino acid makes the difference: the formation of ent-kaurene and 16alpha-hydroxy-ent-kaurane by diterpene synthases in poplar. BMC Plant Biol. 15**,** 262.

Jackson, A. J., Hershey, D. M., Chesnut, T., Xu, M., Peters, R. J., 2014. Biochemical characterization of the castor bean *ent*-kaurene synthase(-like) family supports quantum chemical view of diterpene cyclization. Phytochemistry. 103**,** 13-21.

Jia, M., Peters, R. J., 2016. Extending a single residue switch for abbreviating catalysis in plant ent-kaurene synthases. Frontiers in Plant Science. In press.

Jia, M., Potter, K. C., Peters, R. J., 2016. Extreme promiscuity of a bacterial and a plant diterpene synthase enables combinatorial biosynthesis. Metab. Eng. 37**,** 24-34.

Johnson, S. R., Bhat, W. W., Bibik, J., Turmo, A., Hamberger, B., Genomics Consortium, E. M., Hamberger, B., 2018. A database-driven approach identifies additional diterpene synthase activities in the mint family (Lamiaceae). J Biol Chem.

Kumar, S., Kempinski, C., Zhuang, X., Norris, A., Mafu, S., Zi, J., Bell, S. A., Nybo, S. E., Kinison, S. E., Jiang, Z., Goklany, S., Linscott, K. B., Chen, X., Jia, Q., Brown, S. D.,

Bowman, J. L., Babbitt, P. C., Peters, R. J., Chen, F., Chappell, J., 2016. Molecular Diversity of Terpene Synthases in the Liverwort Marchantia polymorpha. Plant Cell. 28**,** 2632-2650.

Liu, W., Feng, X., Zheng, Y., Huang, C. H., Nakano, C., Hoshino, T., Bogue, S., Ko, T. P., Chen, C. C., Cui, Y., Li, J., Wang, I., Hsu, S. T., Oldfield, E., Guo, R. T., 2014. Structure, function and inhibition of ent-kaurene synthase from Bradyrhizobium japonicum. Sci Rep. 4**,** 6214.

Mafu, S., Fischer, E., Addison, J. B., Riberio Barbosana, I., Zerbe, P., 2016. Substitution of Two Active-Site Residues Alters C9-Hydroxylation in a Class II Diterpene Synthase. Chembiochem : a European journal of chemical biology. 17**,** 2304-2307.

Mafu, S., Hillwig, M. L., Peters, R. J., 2011. A novel labda-7,13e-dien-15-ol-producing bifunctional diterpene synthase from *Selaginella moellendorffii*. ChemBioChem. 12**,** 1984-7.

Mafu, S., Potter, K. C., Hillwig, M. L., Schulte, S., Criswell, J., Peters, R. J., 2015. Efficient heterocyclisation by (di)terpene synthases. Chem Commun (Camb). 51**,** 13485-7.

Mann, F. M., Xu, M., Chen, X., Fulton, D. B., Russell, D. G., Peters, R. J., 2009. Edaxadiene: a new bioactive diterpene from Mycobacterium tuberculosis. Journal of the American Chemical Society. 131**,** 17526-7.

Martin, V. J., Pitera, D. J., Withers, S. T., Newman, J. D., Keasling, J. D., 2003. Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. Nature biotechnology. 21**,** 796-802.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., Olson, A. J., 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of computational chemistry. 30**,** 2785-2791.

Morrone, D., Chambers, J., Lowry, L., Kim, G., Anterola, A., Bender, K., Peters, R. J., 2009. Gibberellin biosynthesis in bacteria: Separate *ent*-copalyl diphosphate and *ent*-kaurene synthases in *Bradyrhizobium japonicum*. FEBS Lett. 583**,** 475-480.

Morrone, D., Hillwig, M. L., Mead, M. E., Lowry, L., Fulton, D. B., Peters, R. J., 2011. Evident and latent plasticity across the rice diterpene synthase family with potential implications for the evolution of diterpenoid metabolism in the cereals. Biochem. J. 435**,** 589-595.

Morrone, D., Jin, Y., Xu, M., Choi, S.-Y., Coates, R. M., Peters, R. J., 2006. An unexpected diterpene cyclase from rice: Functional identification of a stemodene synthase. Arch. Biochem. Biophys. 448**,** 133-140.

Morrone, D., Lowry, L., Determan, M. K., Hershey, D. M., Xu, M., Peters, R. J., 2010. Increasing diterpene yield with a modular metabolic engineering system in E. coli: comparison of MEV and MEP isoprenoid precursor pathway engineering. Appl. Microbiol. Biotechnol. 85**,** 1893-1906.

Nagashima, F., Suzuki, M., Takaoka, S., Asakawa, Y., 2001. Sesqui- and diterpenoids from the Japanese liverwort Jungermannia infusca. Journal of natural products. 64**,** 1309-17.

Nagel, R., Bernholz, C., Vranova, E., Kosuth, J., Bergau, N., Ludwig, S., Wessjohann, L., Gershenzon, J., Tissier, A., Schmidt, A., 2015. Arabidopsis thaliana isoprenyl diphosphate synthases produce the C25 intermediate geranylfarnesyl diphosphate. The Plant journal : for cell and molecular biology. 84**,** 847-59.

Nagel, R., Peters, R. J., 2017. Investigating the Phylogenetic Range of Gibberellin Biosynthesis in Bacteria. Mol Plant Microbe Interact. 30**,** 343-349.

Nakano, C., Hoshino, H., Sato, T., Toyomasu, T., Dairi, T., Sassa, T., 2010. Substrate specificity of the CYC2 enzyme from *Kitasatospora griseola*: production of sclarene, biformene, and

novel bicyclic diterpenes by the enzymatic reactions of labdane-and halimane-type diterpene diphosphates. Tetrahedron Lett. 51**,** 125-128.
Nakano, C., Oshima, M., Kurashima, N., Hoshino, T., 2015. Identification of a new diterpene biosynthetic gene cluster that produces O-methylkolavelool in Herpetosiphon aurantiacus. Chembiochem. 16**,** 772-81.
Noma, M., Suzuki, F., Gamou, K., Kawashima, N., 1982. Two labdane diterpenoids from Nicotiana raimondii. Phytochemistry. 21**,** 395-397.
Pateraki, I., Andersen-Ranberg, J., Hamberger, B., Heskes, A. M., Martens, H. J., Zerbe, P., Bach, S. S., Moller, B. L., Bohlmann, J., Hamberger, B., 2014. Manoyl Oxide (13R), the Biosynthetic Precursor of Forskolin, Is Synthesized in Specialized Root Cork Cells in Coleus forskohlii. Plant Physiol. 164**,** 1222-36.
Pelot, K. A., Chen, R., Hagelthorn, D. M., Young, C. A., Addison, J. B., Muchlinski, A., Tholl, D., Zerbe, P., 2018. Functional Diversity of Diterpene Synthases in the Biofuel Crop Switchgrass. Plant Physiol. 178**,** 54-71.
Pelot, K. A., Mitchell, R., Kwon, M., Hagelthorn, D. M., Wardman, J. F., Chiang, A., Bohlmann, J., Ro, D. K., Zerbe, P., 2017. Biosynthesis of the psychotropic plant diterpene salvinorin A: Discovery and characterization of the Salvia divinorum clerodienyl diphosphate synthase. Plant J. 89**,** 885-897.
Pemberton, T. A., Christianson, D. W., 2016. General base-general acid catalysis by terpenoid cyclases. The Journal of antibiotics. 69**,** 486-93.
Peters, R. J., 2010. Two rings in them all: The labdane-related diterpenoids. Nat. Prod. Rep. 27**,** 1521-1530.
Peters, R. J., Flory, J. E., Jetter, R., Ravn, M. M., Lee, H.-J., Coates, R. M., Croteau, R. B., 2000. Abietadiene synthase from grand fir (*Abies grandis*): Characterization and mechanism of action of the "pseudomature" recombinant enzyme. Biochemistry. 39**,** 15592-15602.
Peters, R. J., Ravn, M. M., Coates, R. M., Croteau, R. B., 2001. Bifunctional abietadiene synthase: Free diffusive transfer of the (+)-copalyl diphosphate intermediate between two distinct active sites. J. Am. Chem. Soc. 123**,** 8974-8978.
Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., Ferrin, T. E., 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 25**,** 1605-12.
Potter, K., Criswell, J., Zi, J., Stubbs, A., Peters, R. J., 2014. Novel product chemistry from mechanistic analysis of ent-copalyl diphosphate synthases from plant hormone biosynthesis. Angewandte Chemie. 53**,** 7198-202.
Potter, K. C., Jia, M., Hong, Y. J., Tantillo, D., Peters, R. J., 2016a. Product Rearrangement from Altering a Single Residue in the Rice syn-Copalyl Diphosphate Synthase. Org Lett. 18**,** 1060-3.
Potter, K. C., Zi, J., Hong, Y. J., Schulte, S., Malchow, B., Tantillo, D. J., Peters, R. J., 2016b. Blocking Deprotonation with Retention of Aromaticity in a Plant ent-Copalyl Diphosphate Synthase Leads to Product Rearrangement. Angewandte Chemie. 55**,** 634-8.
Ramachandran, S., Kota, P., Ding, F., Dokholyan, N. V., 2011. Automated minimization of steric clashes in protein structures. Proteins. 79**,** 261-70.
Sakamoto, T., Miura, K., Itoh, H., Tatsumi, T., Ueguchi-Tanaka, M., Ishiyama, K., Kobayashi, M., Agrawal, G. K., Takeda, S., Abe, K., Miyao, A., Hirochika, H., Kitano, H., Ashikari, M., Matsuoka, M., 2004. An overview of gibberellin metabolism enzyme genes and their related mutants in rice. Plant Physiol. 134**,** 1642-1653.

Sakurai, A., Okamoto, Y., Kokubo, S., CHIDA, a., 1999. Abietane-type Diterpenoids from the Fruit of Vitex rotundifolia L. fil. The Chemical Society of Japan. 3**,** 207-211.
Sali, A., Sanchez, R., Badretdinov, A., 1997. A Program for Protein Structure Modeling Release 4.
Šali, A., Webb, B., Madhusudhan, M., Shen, M.-Y., Dong, G., Martı-Renom, M. A., Eswar, N., Alber, F., Topf, M., Oliva, B., 2014. Azat Badretdinov, Francisco Melo, John P. Overington, and Eric Feyfant email: modeller-care AT salilab. org URL http://salilab. org/modeller.
Sato, T., Yamaga, H., Kashima, S., Murata, Y., Shinada, T., Nakano, C., Hoshino, T., 2013. Identification of novel sesterterpene/triterpene synthase from Bacillus clausii. Chembiochem. 14**,** 822-5.
Schalk, M., Pastore, L., Mirata, M. A., Khim, S., Schouwey, M., Deguerry, F., Pineda, V., Rocci, L., Daviet, L., 2012. Towards a Biosynthetic Route to Sclareol and Amber Odorants. J. Am. Chem. Soc. 134**,** 18900-18903.
Shimane, M., Ueno, Y., Morisaki, K., Oogami, S., Natsume, M., Hayashi, K., Nozaki, H., Kawaide, H., 2014. Molecular evolution of the substrate specificity of ent-kaurene synthases to adapt to gibberellin biosynthesis in land plants. Biochem J. 462**,** 539-46.
Tantillo, D. J., 2017. Importance of Inherent Substrate Reactivity in Enzyme-Promoted Carbocation Cyclization/Rearrangements. Angew Chem Int Ed Engl. 56**,** 10040-10045.
Tawfik, D. S., 2014. Accuracy-rate tradeoffs: how do enzymes meet demands of selectivity and catalytic efficiency? Curr Opin Chem Biol. 21**,** 73-80.
Trott, O., Olson, A. J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 31**,** 455-61.
Whittington, D. A., Wise, M. L., Urbansky, M., Coates, R. M., Croteau, R. B., Christianson, D. W., 2002. Bornyl diphosphate synthase: structure and strategy for carbocation manipulation by a terpenoid cyclase. Proc Natl Acad Sci U S A. 99**,** 15375-15380.
Wilderman, P. R., Xu, M., Jin, Y., Coates, R. M., Peters, R. J., 2004. Identification of *syn*-pimara-7,15-diene synthase reveals functional clustering of terpene synthases involved in rice phytoalexin/allelochemical biosynthesis. Plant Physiol. 135**,** 2098-2105.
Wu, Y., Zhou, K., Toyomasu, T., Sugawara, C., Oku, M., Abe, S., Usui, M., Mitsuhashi, W., Chono, M., Chandler, P. M., Peters, R. J., 2012. Functional characterization of wheat copalyl diphosphate synthases elucidates the early evolution of labdane-related diterpenoid metabolism in the cereals. Phytochemistry. 84**,** 40-46.
Xu, M., Hillwig, M. L., Lane, A. L., Tiernan, M. S., Moore, B. S., Peters, R. J., 2014. Characterization of an orphan diterpenoid biosynthetic operon from Salinispora arenicola. Journal of natural products. 77**,** 2144-7.
Xu, M., Hillwig, M. L., Prisic, S., Coates, R. M., Peters, R. J., 2004. Functional identification of rice syn-copalyl diphosphate synthase and its role in initiating biosynthesis of diterpenoid phytoalexin/allelopathic natural products. The Plant journal : for cell and molecular biology. 39**,** 309-18.
Xu, M., Jia, M., Hong, Y. J., Yin, X., Tantillo, D. J., Proteau, P. J., Peters, R. J., 2018. Premutilin Synthase: Ring Rearrangement by a Class II Diterpene Cyclase. Org Lett. 20**,** 1200-1202.
Xu, M., Wilderman, P. R., Morrone, D., Xu, J., Roy, A., Margis-Pinheiro, M., Upadhyaya, N., Coates, R. M., Peters, R. J., 2007. Functional characterization of the rice kaurene synthase-like gene family. Phytochemistry. 68**,** 312-326.
Yamada, Y., Komatsu, M., Ikeda, H., 2016. Chemical diversity of labdane-type bicyclic diterpene

biosynthesis in Actinomycetales microorganisms. The Journal of antibiotics. 69**,** 515-23.
Yamaguchi, S., Sun, T., Kawaide, H., Kamiya, Y., 1998. The GA2 locus of *Arabidopsis thaliana* encodes *ent*-kaurene synthase of gibberellin biosynthesis. Plant Physiol. 116**,** 1271-1278.
Zerbe, P., Chiang, A., Dullat, H., O'Neil-Johnson, M., Starks, C., Hamberger, B., Bohlmann, J., 2014. Diterpene synthases of the biosynthetic system of medicinally active diterpenoids in Marrubium vulgare. Plant J. 79**,** 914-27.
Zerbe, P., Chiang, A., Yuen, M., Hamberger, B., Draper, J. A., Britton, R., Bohlmann, J., 2012. Bifunctional *cis*-abienol synthase from *Abies balsamea* discovered by transcriptome sequencing and its implications for diterpenoid fragrance production. J. Biol. Chem. 287**,** 12121-31.
Zerbe, P., Hamberger, B., Yuen, M. M., Chiang, A., Sandhu, H. K., Madilao, L. L., Nguyen, A., Hamberger, B., Bach, S. S., Bohlmann, J., 2013. Gene discovery of modular diterpene metabolism in nonmodel systems. Plant Physiol. 162**,** 1073-91.
Zhou, K., Xu, M., Tiernan, M. S., Xie, Q., Toyomasu, T., Sugawara, C., Oku, M., Usui, M., Mitsuhashi, W., Chono, M., Chandler, P. M., Peters, R. J., 2012. Functional characterization of wheat ent-kaurene(-like) synthases indicates continuing evolution of labdane-related diterpenoid metabolism in the cereals. Phytochemistry. 84**,** 47-55.
Zi, J., Mafu, S., Peters, R. J., 2014a. To Gibberellins and Beyond! Surveying the Evolution of (Di)Terpenoid Metabolism. Annu. Rev. Plant Biol. 65**,** 259-286.
Zi, J., Matsuba, Y., Hong, Y., Jackson, A., Pichersky, E., Tantillo, D. J., Peters, R. J., 2014b. Biosynthesis of lycosantalonol, a *cis*-prenyl derived diterpenoid. J. Am. Chem. Soc. 136**,** 16951-16953.

Tables

Table 1. Potential substrates and the relevant synthases used in this study.

No.[a]	Name[b]	DTC/IDS[c]	Origin	Reference
1	GGPP	GGPPS	*Abies grandis*	(Burke and Croteau, 2002)
2	NNPP	NNPPS	*Solanum lycopersicum*	(Zi et al., 2014b)
3	FPP	FPPS	*Saccharomyces cerevisiae*	(Martin et al., 2003)
4	GFPP	GFPPS	*Arabidopsis thaliana*	(Nagel et al., 2015)
5	copalyl diphosphate (CPP)	AgAS:D621A	*Abies grandis*	(Peters et al., 2001)
6	*ent*-CPP	An2/ZmCPS2	*Zea mays*	(Harris et al., 2005)
7	*syn*-CPP	OsCPS4	*Oryza sativa*	(Xu et al., 2004)
8	7-*endo*-CPP	SmCPS/KSL1:D501A/ D505A	*Selaginella moellendorffii*	(Mafu et al., 2011)
9	*ent*-7-*endo*-CPP	SdCPS2:C359F/W360S	*Salvia divinorum*	(Pelot et al., 2017)
10	8α-hydroxy-CPP	NgCLS	*Nicotiana glutinosa*	(Criswell et al., 2012)
11	8β-hydroxy-*ent*-CPP	AtCPS:H263A	*Arabidopsis thaliana*	(Potter et al., 2014)
12	peregrinol diphosphate (9α-hydroxy-CPP)	MvCPS1	*Marrubium vulgare*	(Zerbe et al., 2014)
13	kolavenyl diphosphate (KPP)	Haur_2145	*Herpetosiphon aurantiacus*	(Nakano et al., 2015)
14	*ent*-KPP	AtCPS:H263Y	*Arabidopsis thaliana*	(Potter et al., 2016b)
15	terpentedienyl diphosphate (*syn*-KPP)	KgTPS	*Kitasatospora griseola*	(Nakano et al., 2010)
16	tuberculosinyl diphosphate (halimadienyl diphosphate, HPP)	MtHPS	*Mycobacterium tuberculosis*	(Mann et al., 2009)
17	*syn*-HPP	OsCPS4:H561D	*Oryza sativa*	(Potter et al., 2016a)
18	*syn*-halima-5(10),13*E*-dienyl diphosphate	MvCPS1:W323F/F505Y	*Marrubium vulgare*	(Mafu et al., 2016)
19	mutildienyl diphosphate	CpPS:D649L	*Clitopilus passeckerianus*	(Xu et al., 2018)

[a]Numbering as defined in the text.

[b]Previously assigned common names, with semi-systematic names (as defined in the text) also given where necessary.

[c]Full names of these synthases can be found in the abbreviation list.

Table 2. Relative reactivity of the ten DTSs and 19 potential substrates examined here.

DTS[a] / substrate	SsSS[b]	AbCAS	SmMS	OsKS	AtKS	KgTS[b]	ScLS	SaPS	BjKS	EtKS
1	+++	-	-	-	-	+++	++	++	+	++
2	++	-	-	-	-	+++	+	-	-	-
3	+	-	-	-	-	+	+	-	+	+
4	-	-	-	-	-	++	-	-	-	-
5	+++	+	**+++**	-	-	+++	**+++**	**+++**	++	+++
6	+++	-	-	**+++**	**+++**	+++	+++	++	**+++**	**+++**
7	+++	-	+++	-	-	+++	+++	+++	+++	+
8	+++	-	-	-	-	+++	+++	+	+++	+++
9	+++	-	-	-	+	+++	+++	++	+++	+++
10	**+++**	**+++**	+++	-	-	+++	+++	+++	+++	+++
11	+++	-	-	+++	+++	+++	+++	+++	+++	+++
12	+++	-	-	-	-	+++	+++	++	+++	+++
13	+++	-	-	-	-	+++	+++	-	-	-
14	+++	-	-	-	-	+++	+++	-	+++	+++
15	+++	-	-	-	-	**+++**	+++	-	-	-
16	+++	-	-	-	-	+++	+++	-	-	-
17	+++	-	-	-	-	+++	+++	+	+	+++
18	+++	-	-	-	-	+++	+++	+	+	++
19	-	-	-	-	-	+++	+	-	-	-

"-" indicates substrate conversion percentage (P) < 10%; "+" 10% ≤ P < 40%; "++" 40% ≤ P < 70%; "+++" 70% ≤ P ≤ 100%; specific substrate conversion percentage values can be found in Fig. S1 and Table S7. Native substrate indicated by **bold** red text.

[a]Plant DTSs indicated by green text and bacterial by blue text.

[b]The substrate conversion percentage values for SsSS and KgTS with substrates **5-8** and **10-17** were obtained from a previous report (Jia et al., 2016).

Table 3. Overview of the DTS products obtained from this study[a].

No.[b]	Semi-systematic name[c]	with stereo-definitions	Common name[d]	Identification[e]
20			β-farnesene	Prev. (Hamano et al., 2002)
21			(*Z*)-α-farnesene	Prev. (Hamano et al., 2002)
22			*(E)-α*-farnesene	Prev. (Hamano et al., 2002)
23			(*E*)-nerolidol	Comp. (commercial std.)
24			β-geranylfarnesene	Comp. (Sato et al., 2013)
25			(*Z*)-α-geranylfarnesene	This study
26			(*E*)-α-geranylfarnesene	This study
27			β-nerylmyrcene	This study
28			β-springene	Prev. (Jia et al., 2016)
29	*ent*-labda-7,13(16),14-triene	(9*S*,10*R*)-labda-7, 13(16),14-triene		This study
30	*ent*-labda-7,14-dien-13-ol	(9*S*,10*R*)-labda-7,14-dien-13-ol		This study
31	*syn*-halima-5(10), 13(16),14-triene	(8*R*,9*R*)-halima-5(10), 13(16),14-triene		This study
32	*syn*-halima-5(10),14-dien-13-ol	(8*R*,9*R*)-halima-5(10), 14-dien-13-ol		This study
33	mutil-4(18),13(16),14-triene[f]			This study
34	labda-12*Z*,14-dien-8α-ol,	(8*R*,9*R*,10*S*)-labda-12Z,14-dien-8α-ol	*cis*-abienol	Prev. (Zerbe et al., 2012)
35	labda-8(17),12*E*,14-triene	(9*S*,10*S*)-labda-8(17), 12*E*,14-triene		Prev. (Ikeda et al., 2016; Yamada et al., 2016)
36	labda-7,12*E*,14-triene	(9*S*,10*S*)-labda-7,12*E*, 14-triene		Prev. (Ikeda et al., 2016; Yamada et al., 2016)
37			(*E*)-α-springene	Comp. (Jia et al., 2016)
38	labda-8(17),13(16),14-triene	(9*S*,10*S*)-labda-8(17), 13(16),14-triene	sclarene	Comp. (Jia et al., 2016)
39	labda-8(17),12*Z*,14-triene	(9*S*,10*S*)-labda-8(17), 12Z,14-triene	(*Z*)-biformene	This study
40	*ent*-labda-8(17), 13(16),14-triene	(9*R*,10R)-labda-8(17), 13(16),14-triene	*ent*-sclarene	This study
41		(8*S*,9*R*,10*R*,13*R*)-kaur-16-ene	*ent*-kaurene	Comp. (Morrone et al., 2009)
42	*syn*-labda-8(17), 13(16),14-triene	(9*R*,10*S*)-labda-8(17),13(16),14-triene	griseolaene	Comp. (Jia et al., 2016)
43	*syn*-labda-8(17),14-dien-13-ol	(9*R*,10*S*,13*S*)-labda-8(17),14-dien-13-ol	vitexifolin A	Comp. (Jia et al., 2016)
44	*ent*-labda-7,12*E*,14-triene	(9*R*,10*R*)-labda-7,12*E*, 14-triene		This study
45	*ent*-labda-8,13*R*-epoxy-14-ene	(8*S*,9*S*,10*R*,13*R*)-labda-8,13-epoxy-14-ene	*ent*-13-epi-manoyl oxide	Comp. (Mafu et al., 2015)
46	cleroda-3,14-dien-13-ol	(5*S*,8*S*,9*R*,10*S*)-cleroda-3,14-dien-13-ol	kolavelool	Comp. (Jia et al., 2016)
47	labda-8,13*R*-epoxy-14-ene	(8*R*,9*R*,10*S*,13*R*)-labda-8,13-epoxy-14-ene	manoyl oxide	Comp. (Mafu et al., 2015)
48	labda-8,13*S*-epoxy-14-ene	(8*R*,9*R*,10*S*,13*S*)-labda-8,13-epoxy-14-ene	13-*epi*-manoyl oxide	Comp. (Mafu et al., 2015)
49	labda-12*E*,14-dien-8α-ol	(8*R*,9*R*,10*S*)-labda-12*E*,14-dien-8α-ol	*trans*-abienol	This study
50	*syn*-labda-9,13*S*-epoxy-14-ene	(8*R*,9*R*,10*S*,13*S*)-labda-9,13-epoxy-14-ene		This study

51	*syn*-labda-9,13*R*-epoxy-14-ene	(8*R*,9*R*,10*S*,13*R*)-labda-9,13-epoxy-14-ene		This study
52	*ent*-cleroda-3,12*E*,14-triene	(5*R*,8*R*,9*S*,10*R*)-cleroda-3,12*E*,14-triene		This study
53	*syn*-cleroda-3,12*E*,14-triene	(5*S*,8*R*,9*R*,10*S*)-cleroda-3,12*E*,14-triene		This study
54	*syn*-cleroda-3,13(16),14-triene	(5*S*,8*R*,9*R*,10*S*)-cleroda-3,13(16),14-triene	terpentetriene	Comp. (Jia et al., 2016)
55	halima-5,12*E*,14-triene	(8*S*,9*R*,10*S*)-halima-5,12*E*,14-triene		This study
56	*syn*-halima-5, 13(16),14-triene	(8*R*,9*R*,10*S*)-halima-5,13(16),14-triene		Comp. (Jia et al., 2016)
57	*syn*-halima-5,12*E*, 14-triene	(8*R*,9*R*,10*S*)-halima-5,12*E*,14-triene		This study
58	*syn*-halima-5(10), 12*E*,14-triene	(8*R*,9*R*)-halima-5(10), 12*E*,14-triene		This study
59	abieta-8,12-diene	(10*S*)-abieta-8,12-diene	miltiradiene	Comp. (Gao et al., 2009)
60	isopimara-8,15-diene	(10*S*,13*S*)-isopimara-8,15-diene		Comp. (Xu et al., 2014)
61			geranyllinalool	Comp. (Jia et al., 2016)
62	labda-8(17),14-dien-13-ol	(9*S*,10*S*,13*R*)-labda-8(17),14-dien-13-ol	manool	Comp. (Jia et al., 2016)
63	pimara-7,15-diene	(9*S*,10*S*,13*R*)-pimara-7,15-diene		This study
64	*ent*-labda-8(17),14-dien-13-ol	(9*R*,10*R*,13*S*)-labda-8(17),14-dien-13-ol	*ent*-manool	Comp. (Jia et al., 2016)
65	*ent*-pimara-8(14),15-diene	(9*R*,10*R*,13*R*)-pimara-8(14),15-diene		Comp. (Jia and Peters, 2016)
66	labda-7,14-dien-13-ol	*(9R,10S)*-labda-7,14-dien-13-ol		Comp. (Jia et al., 2016)
67	*syn*-labda-14-en-9,13-diol	(8*R*,9*R*,10*S*)-labda-14-en-9,13-diol	viteagnusin D	Comp. (Jia et al., 2016)
68	*syn*-labda-13(16),14-dien-9-ol	(8*R*,9*R*,10*S*)-labda-13(16),14-dien-9-ol		Comp. (Jia et al., 2016)
69	*ent*-cleroda-3,14-dien-13-ol	(5*R*,8*R*,9*S*,10*R*)-cleroda-3,14-dien-13-ol		Comp. (Jia et al., 2016)
70	*ent*-cleroda-3,13(16),14-triene	(5*R*,8*R*,9*S*,10*R*)-cleroda-3,13(16),14-triene		Comp. (Jia et al., 2016)
71	*syn*-halima-5,14-dien-13-ol	(8*R*,9*R*,10*S*)-halima-5,14-dien-13-ol		Comp.. (Jia et al., 2016)
72	*syn*-pimara-7,15-diene	(9*R*,10*S*,13*R*)-pimara-7,15-diene		Comp. (Wilderman et al., 2004)
73	*syn*-stemodene	(9*R*,10*S*,13*S*)-stemod-13(17)-ene		Comp. (Morrone et al., 2006)
74	*ent*-labda-8,13*S*-epoxy-14-ene	(8*S*,9*S*,10*R*,13*S*)-labda-8,13-epoxy-14-ene	*ent*-manoyl oxide	Comp. (Mafu et al., 2015)
75	*syn*-abieta-9(11),12-diene	(8*S*,10*S*)-abieta-9(11), 12-diene		This study
76	*syn*-pimara-9(11),15-diene	(10*S*,13*R*)-pimara-9(11),15-diene		This study

[a]Relevant DTS and substrate can be found in Table S18.

[b]Products are numbered as defined in the text.
[c]Semi-systematic names are designated for LRD derivatives by final backbone and stereochemically by the configuration of carbons 9 and 10 in the initially formed decalin bicycle (see Fig. 2), with complete stereodefinition also given.
[d]Common names are those previously reported.
[e]Products were identified based on either previous (Prev.) reports for these enzymes or GC-MS based comparison ('Comp.') to other previously reported DTS products (with accompanying reference(s) in these cases), or were determined in 'This study' by NMR based structural analysis or comparison to such a characterized enantiomer).
[f]Numbering here differs from that used in the original report (Xu et al., 2018), to maintain consistency with that of the other LRD products.

Figures & Legends

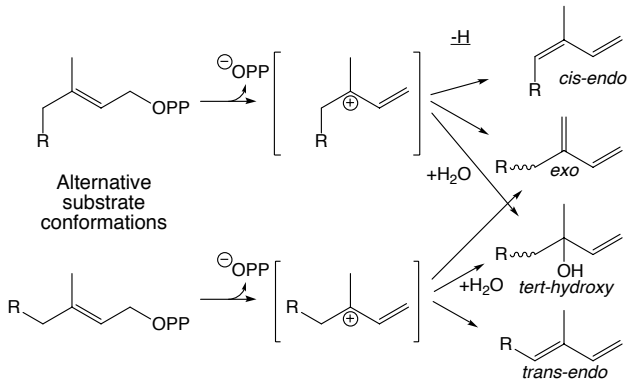


Fig. 1. Simple lysis (of the allylic diphosphate ester) reactions catalyzed by TPSs with variable length isoprenyl precursors (R = one or more isopentenyl units). Direct deprotonation of the initially formed tertiary carbocation intermediate leads to the depicted range of products, with configuration of the newly formed internal *endo* C=C presumably dependent on initial substrate conformation (as shown).

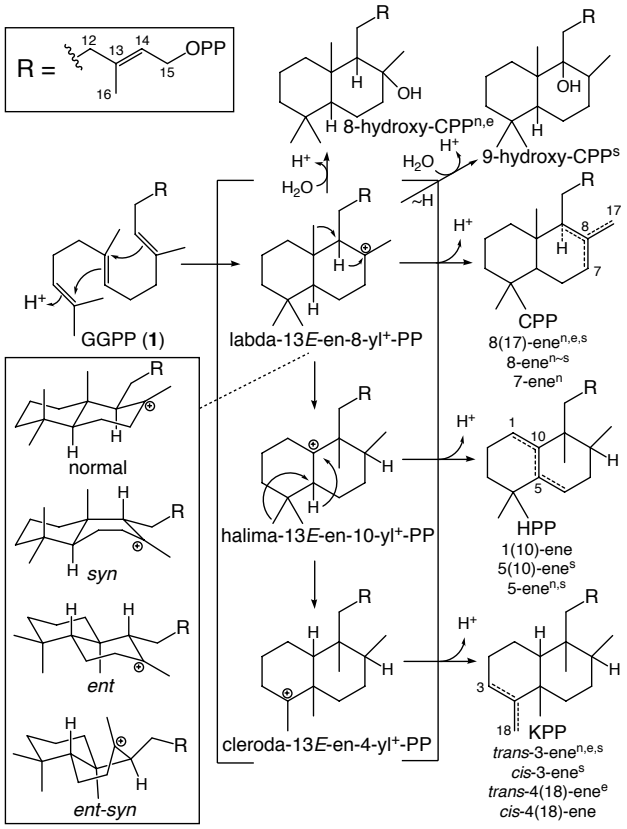


Fig. 2. Basic DTC products from bicyclization and subsequent rearrangement (PP=diphosphate). Also shown are known stereoisomers for the initial bicycle, with derived products for which DTCs are known indicated by superscript ([n]normal, [e]*ent*, [s]*syn*, no *ent-syn* have yet been identified).

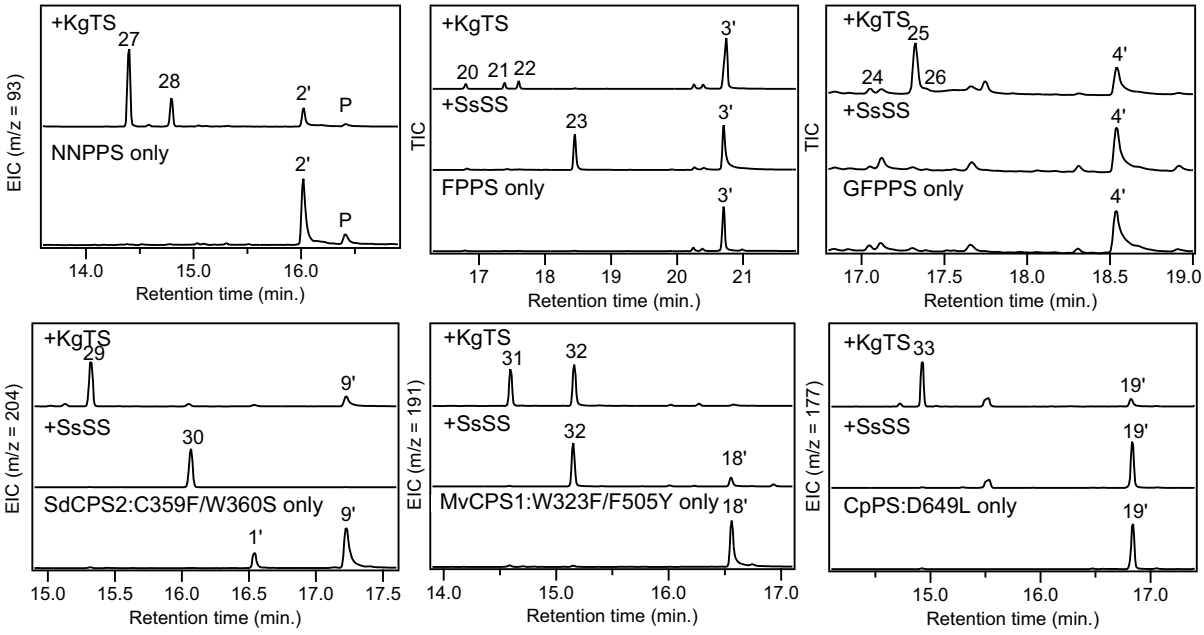


Fig. 3. Extended promiscuity of SsSS and, particularly, KgTS. GC–MS chromatograms (either total or indicated extracted ion count, TIC or EIC, respectively) of extracts from *E. coli* engineered for production of a potential substrate by introduction of either the combination of a relevant DTC and GGPPS, or an alternative IDS only (Table 1), along with KgTS or SsSS, as indicated (peaks are labeled by compound numbers, as described in the text, with those corresponding to the dephosphorylated substrate derivatives indicated by prime' notation).

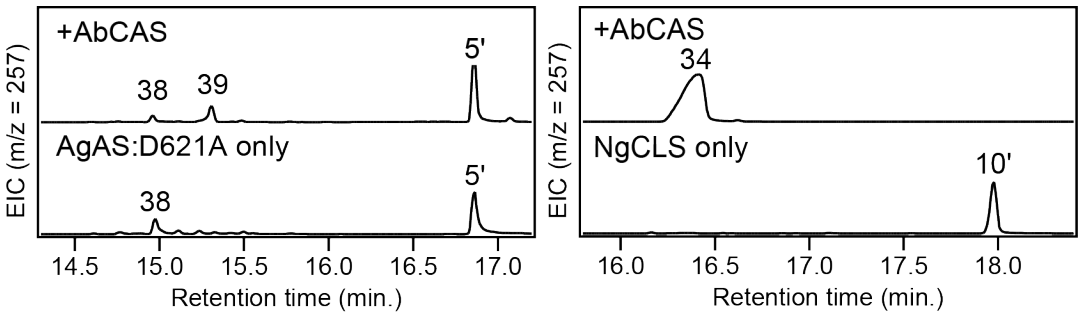


Fig. 4. Limited promiscuity of AbCAS. GC–MS chromatograms of extracts from *E. coli* engineered for production of the indicated potential substrate by co-expression of GGPPS and a relevant DTC (Table 1), along with AbCAS, as indicated (peaks are labeled by compound numbers, as described in the text, with those corresponding to the dephosphorylated substrate derivatives indicated by prime' notation). Only the two examples exhibiting conversion to product(s) are shown, no turnover was observed with all other potential substrates.

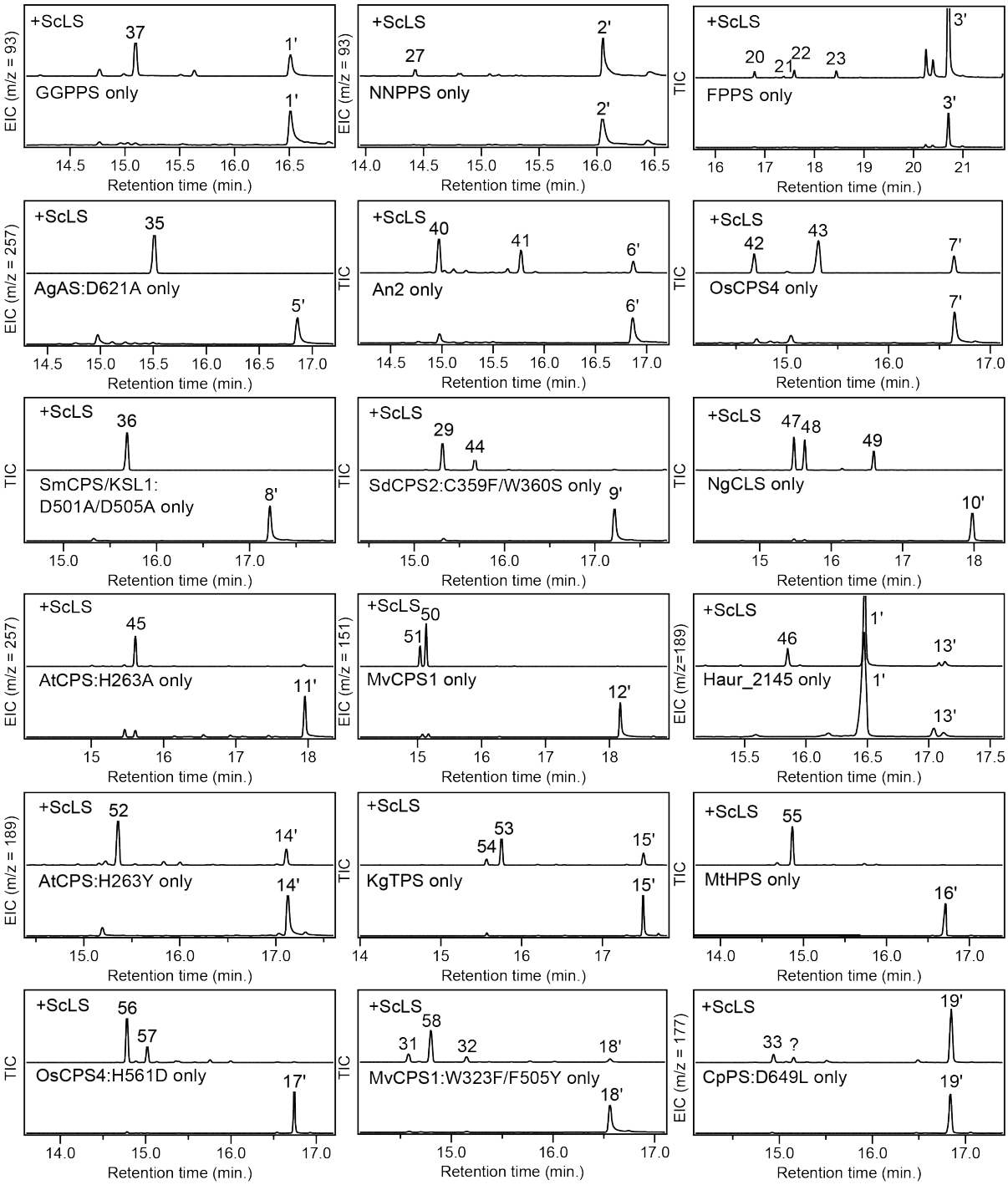


Fig. 5. Extreme promiscuity of ScLS. GC–MS chromatograms of extracts from *E. coli* engineered for production of a potential substrate by introduction of either the combination of a relevant DTC and GGPPS, or an IDS only (Table 1), along with ScLS, as indicated (peaks are labeled by compound numbers, as described in the text, with those corresponding to the dephosphorylated substrate derivatives indicated by prime' notation).

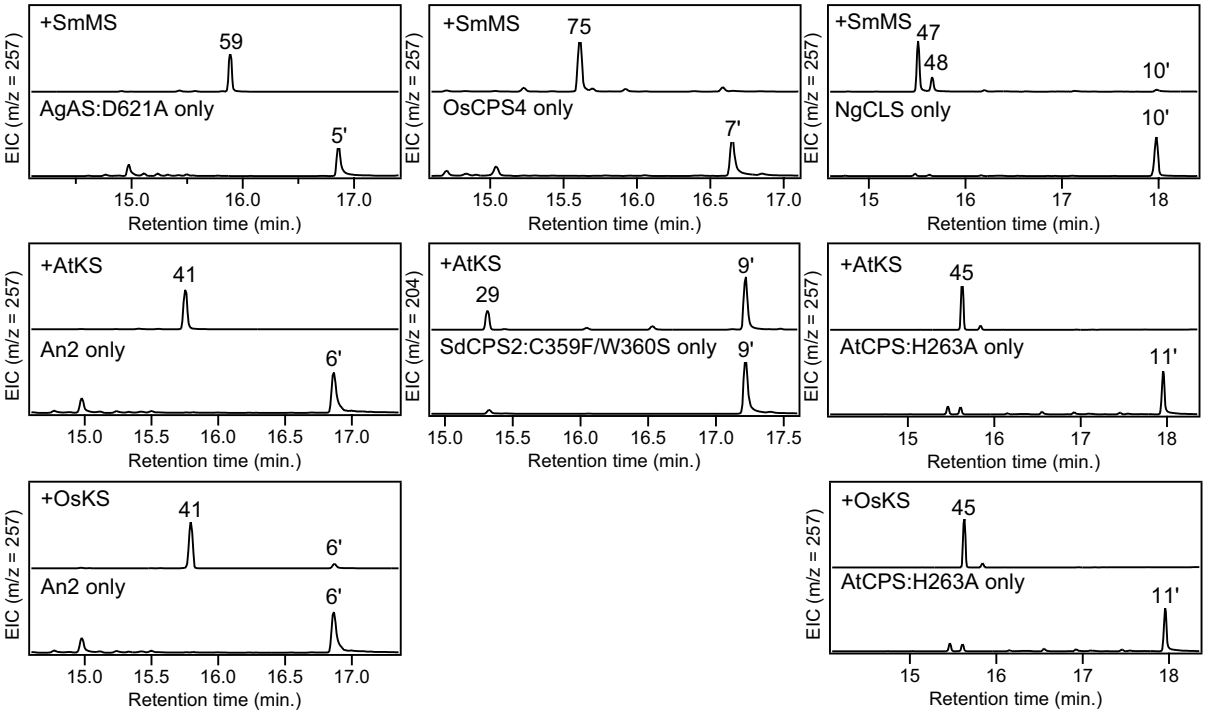


Fig. 6. Limited promiscuity of plant DTSs. GC–MS chromatograms of extracts from *E. coli* engineered for production of the indicated potential substrate by co-expression of GGPPS and a relevant DTC (Table 1), along with a plant DTS, as indicated (peaks are labeled by compound numbers, as described in the text, with those corresponding to the dephosphorylated substrate derivatives indicated by prime' notation). Only combinations exhibiting conversion to product(s) are shown, no turnover was otherwise observed.

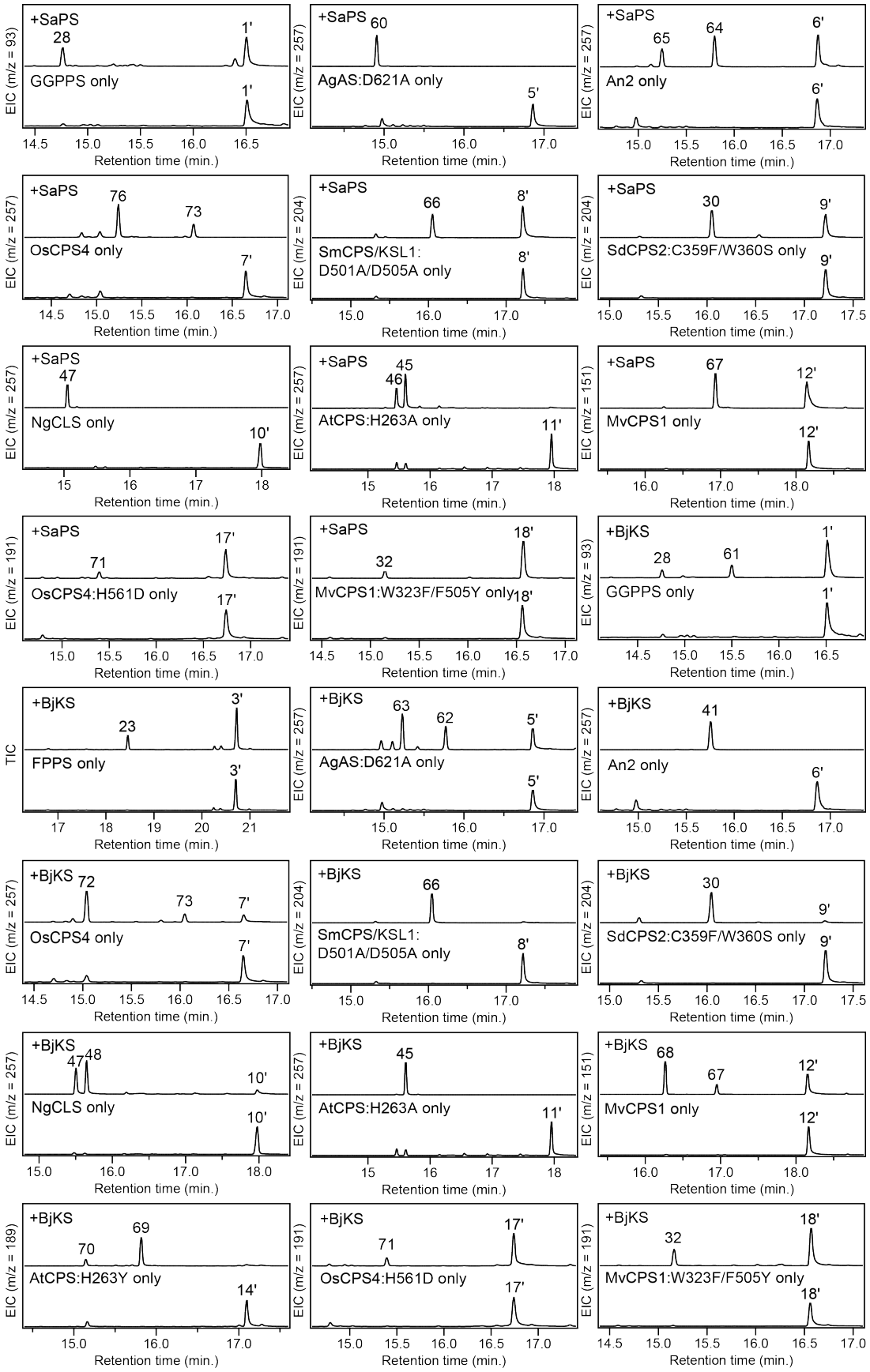

+SaPS
GGPPS only
+SaPS
AgAS:D621A only
+SaPS
An2 only
+SaPS
OsCPS4 only
+SaPS
SmCPS/KSL1:
D501A/D505A only
+SaPS
SdCPS2:C359F/W360S only
+SaPS
NgCLS only
+SaPS
AtCPS:H263A only
+SaPS
MvCPS1 only
+SaPS
OsCPS4:H561D only
+SaPS
MvCPS1:W323F/F505Y only
+BjKS
GGPPS only
+BjKS
FPPS only
+BjKS
AgAS:D621A only
+BjKS
An2 only
+BjKS
OsCPS4 only
+BjKS
SmCPS/KSL1:
D501A/D505A only
+BjKS
SdCPS2:C359F/W360S only
+BjKS
NgCLS only
+BjKS
AtCPS:H263A only
+BjKS
MvCPS1 only
+BjKS
AtCPS:H263Y only
+BjKS
OsCPS4:H561D only
+BjKS
MvCPS1:W323F/F505Y only
Retention time (min.)

Fig. 7. Extreme promiscuity of bacterial DTSs. GC–MS chromatograms of extracts from *E. coli* engineered for production of the indicated potential substrate by co-expression of GGPPS and a relevant DTC (or GGPPS or FPPS alone), along with a bacterial DTS, as indicated (peaks are labeled by compound numbers, as described in the text, with those corresponding to the dephosphorylated substrate derivatives indicated by prime' notation). Only combinations exhibiting conversion to product(s) are shown, no turnover was otherwise observed.

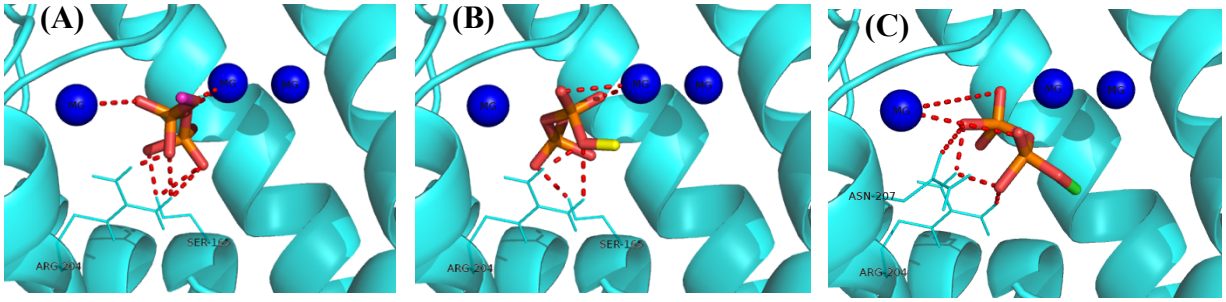


Fig. 8. Molecular docking study of BjKS with three representative substrates highlighting diphosphate positioning (phosphorus in orange and oxygen in red), with the bicyclic olefin substituent simply represented by a methyl group for ease of visualization – *ent*-CPP (**6**) in purple (A), *syn*-KPP (**15**) in yellow (B) and *syn*-HPP (**17**) in green (C). Residues involved in polar contacts with the phosphate group are shown as aqua-blue lines, Mg^{2+} ions shown as dark blue spheres, and polar contacts shown as red dotted lines.

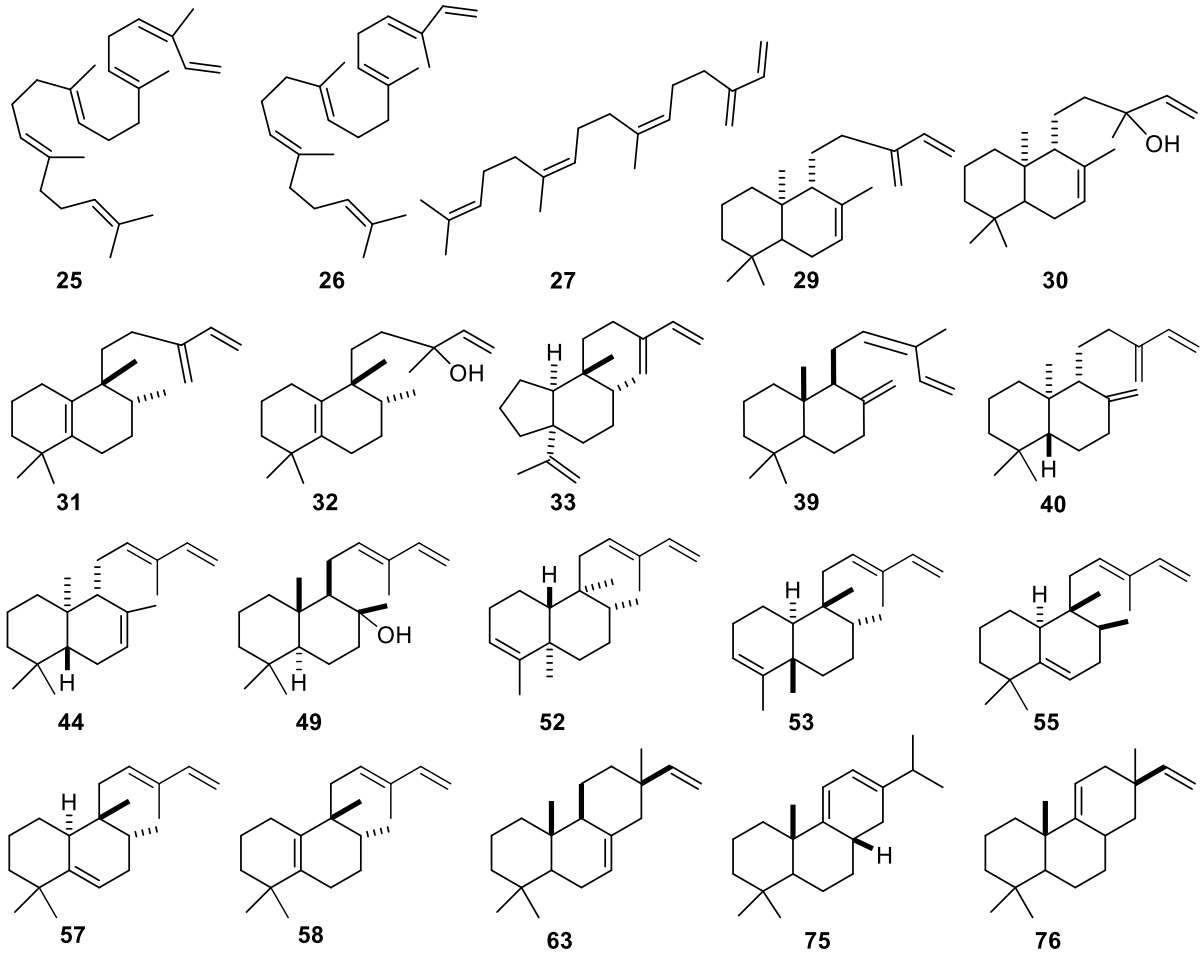


Fig. 9. Chemical structures of newly enzymatic products discovered in this study (numbering as defined in the text).

Supplementary Information for:

Combinatorial biosynthesis and the basis for promiscuity in class I diterpene synthases

Meirong Jia,[a] Sambit K. Mishra,[b] Samuel Tufts, [a] Robert Jernigen,[a,b] Reuben J. Peters[a,]*

[a]Roy J. Carver Department of Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames, IA 50011, USA

[b]LH Baker Center for Bioinformatics and Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA 50011, USA

*Corresponding author: e-mail address: rjpeters@iastate.edu

Contents

Table S1. Recently identified DTCs (not included in this study).

DTC	Product[a]	Origin	Reference
PcCPS1	*ent*-labda-8,13*E*-dienyl diphosphate,	*Pogostemon cablin*	(Johnson, *et al.*, 2018)
PvKPS1:F251V	*cis*-kolava-3,13*E*-dienyl diphosphate	*Panicum virgatum*	(Pelot, *et al.*, 2019)
ArKPS	*ent*-kolava-4(18),13*E*-dienyl diphosphate	*Ajuga reptans*	(Johnson, *et al.*, 2018)
PvKPS1	*cis*-*syn*-kolava-3,13*E*-dienyl diphosphate	*Panicum virgatum*	(Pelot, *et al.*, 2018)

[a]"*cis*" refers to the unusual *cis*-decalin bridgehead (C5,10) configuration.

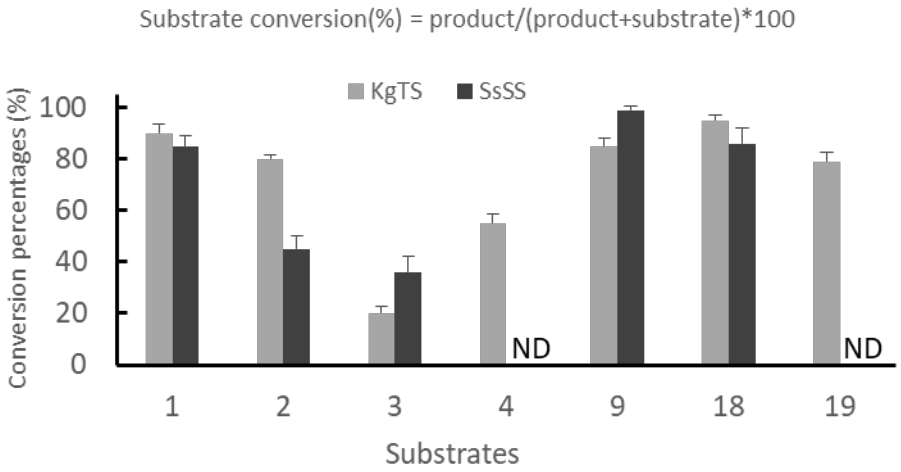


Fig. S1. Substrate conversion percentages of KgTS and SsSS upon co-expression in *E. coli* also engineered to produce the indicated potential substrate. Values are given as Mean ± SD from triplicate measurements. "ND" means activity was not detected.

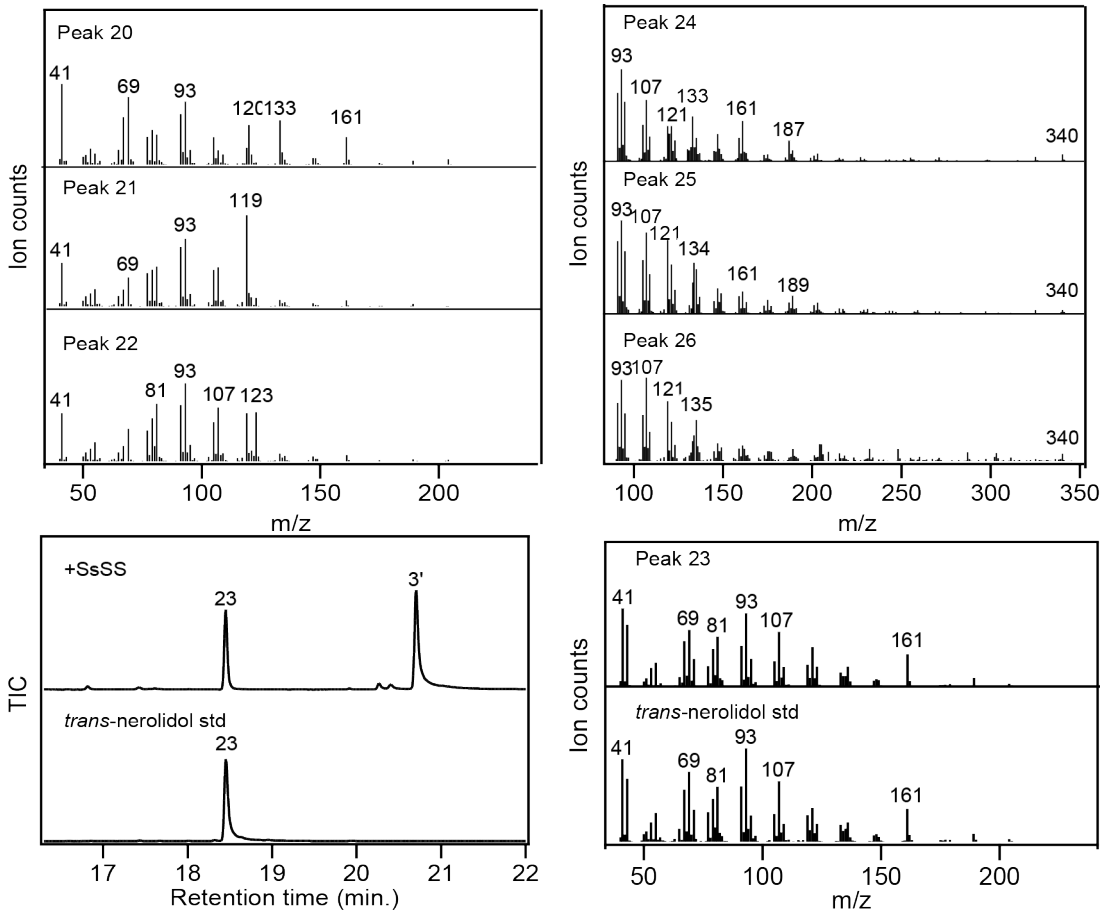


Fig. S2. Product identification via comparison of both RT and MS from GC-MS analyses. Products **20-22** were verified by MS comparison to that previously reported for the same combinations of KgTS and substrate *in vitro* (Nakano*, et al.*, 2010). Product **23** was identified by comparison to a commercially obtained *trans*-nerolidol standard. Product **24** was tentatively identified by MS comparison to that previously reported (Huang, et al., 2017, Sato, et al., 2013), while **25** & **26** are even more tentatively assigned as based on their relative RT and MS to those observed with the shorter **21** & **22**. For general reference the MS are presented here.

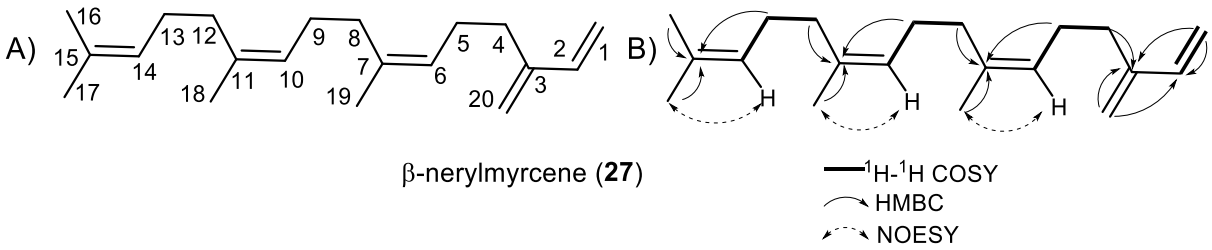


Fig. S3. The major product of KgTS upon co-expression with the isoprenyl diphosphate synthase producing **2** (NNPPS). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-^{1}H COSY correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure.

Table S2. ^{1}H and ^{13}C NMR assignments for compound **27**, β-nerylmyrcene (solvent $CDCl_3$).

Position	β-nerylmyrcene (**27**)	
	δ_H	δ_C
1a	5.21 (1H, d, J = 17.8 Hz)	113.3
b	5.03 (1H, d, J = 10.9 Hz)	
2	6.35 (1H, dd, J = 17.8 Hz, 10.9 Hz)	139.2
3		146.4
4	2.20 (2H, m)	31.9
5	2.17 (2H, m)	26.7
6	5.15 (1H, t, J = 6.9 Hz)	125.1
7		135.7
8	2.02 (2H, m)	32.2
9	2.04 (2H, m)	26.6
10	5.11 (1H, m)	125.2
11		135.6
12	2.02 (2H, m)	32.5
13	2.04 (2H, m)	26.9
14	5.10 (1H, m)	124.6
15		131.8
16	1.59 (3H, s)	17.9
17	1.66 (3H, s)	25.9
18	1.66 (3H, s)	23.62
19	1.68 (3H, s)	23.64
20a	4.99 (1H, s)	115.9
b	4.97 (1H, s)	

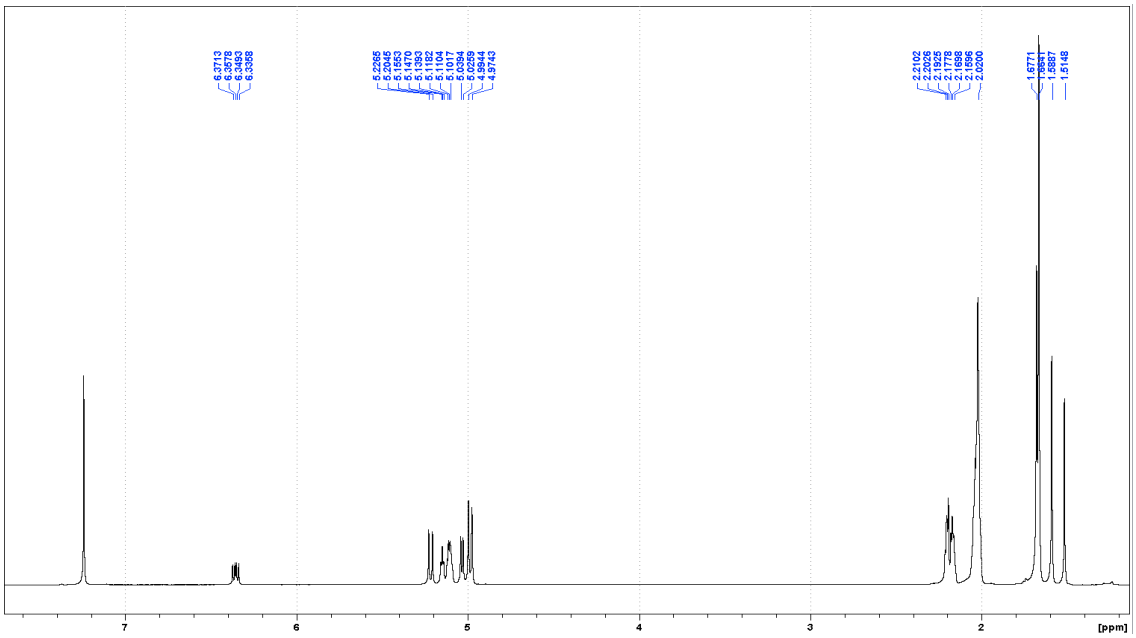


Fig. S4. (A) ^{1}H Spectrum of β-nerylmyrcene (**27**)

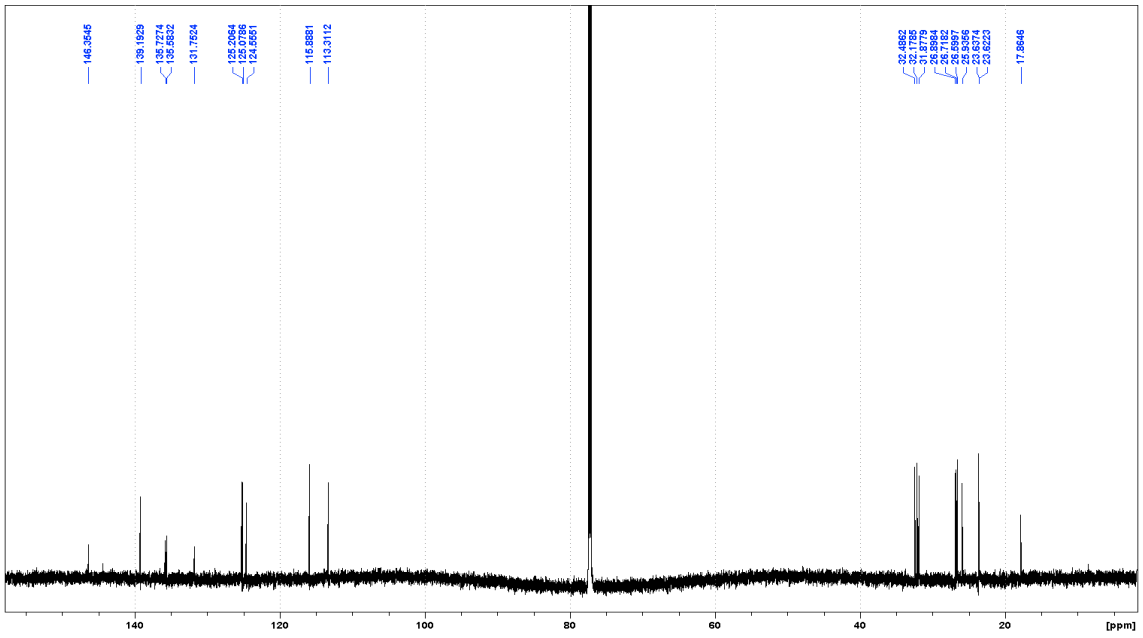


Fig. S4. (B) ^{13}C Spectrum of β-nerylmyrcene (**27**)

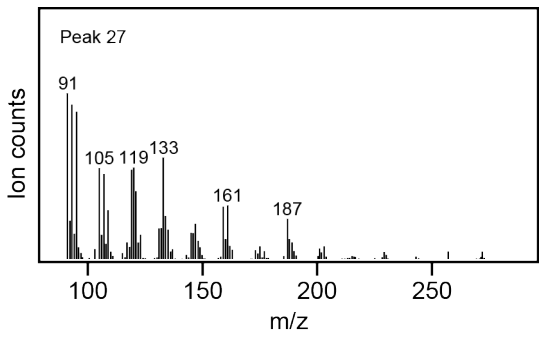


Fig. S5. MS of β-nerylmyrcene (**27**)

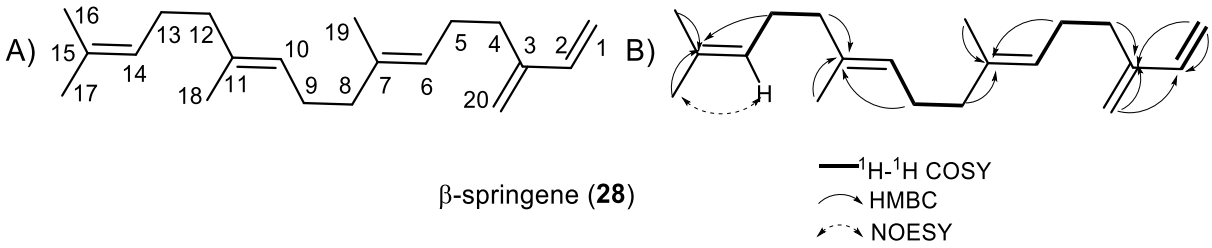


Fig. S6. The minor product of KgTS upon co-expression with the isoprenyl diphosphate synthase producing **2** (NNPPS). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-^{1}H COSY correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure.

Table S3. ^{1}H and ^{13}C NMR assignments for compound **28**, β-springene (solvent $CDCl_3$).

Position	β-springene (**28**)	
	δ_H	δ_C
1a	5.23 (1H, d, J = 17.6 Hz)	113.3
b	5.04 (1H, d, J = 10.9 Hz)	
2	6.36 (1H, dd, J = 17.6 Hz, 10.9 Hz)	139.2
3		146.4
4	2.20 (2H, m)	31.6
5	2.17 (2H, m)	26.83
6	5.15 (1H, t, J = 6.7 Hz)	124.4
7		135.6
8	1.98 (2H, m)	40
9	2.06 (2H, m)	26.85
10	5.10 (1H, m)	124.2
11		135.2
12	1.96 (2H, m)	39.9
13	2.04 (2H, m)	27.0
14	5.08 (1H, m)	124.6
15		131.5
16	1.58 (3H, s)	17.9
17	1.66 (3H, s)	25.9
18	1.58 (3H, s)	16.2
19	1.59 (3H, s)	16.3
20a	4.99 (1H, s)	115.9
b	4.98 (1H, s)	

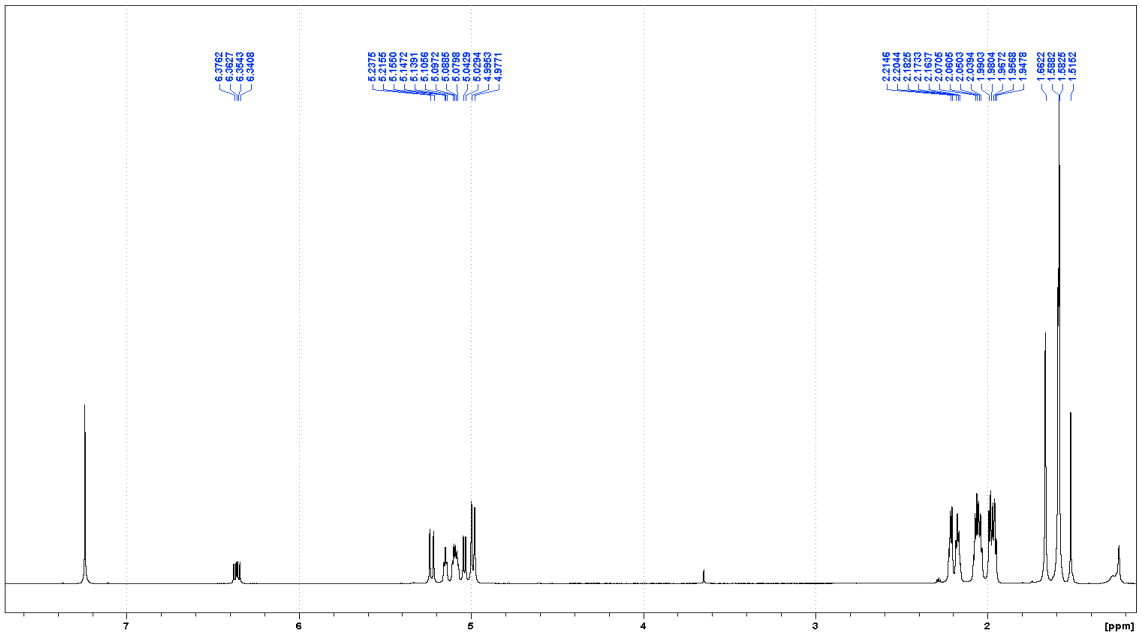


Fig. S7. (A) ^{1}H Spectrum of β-springene (**28**)

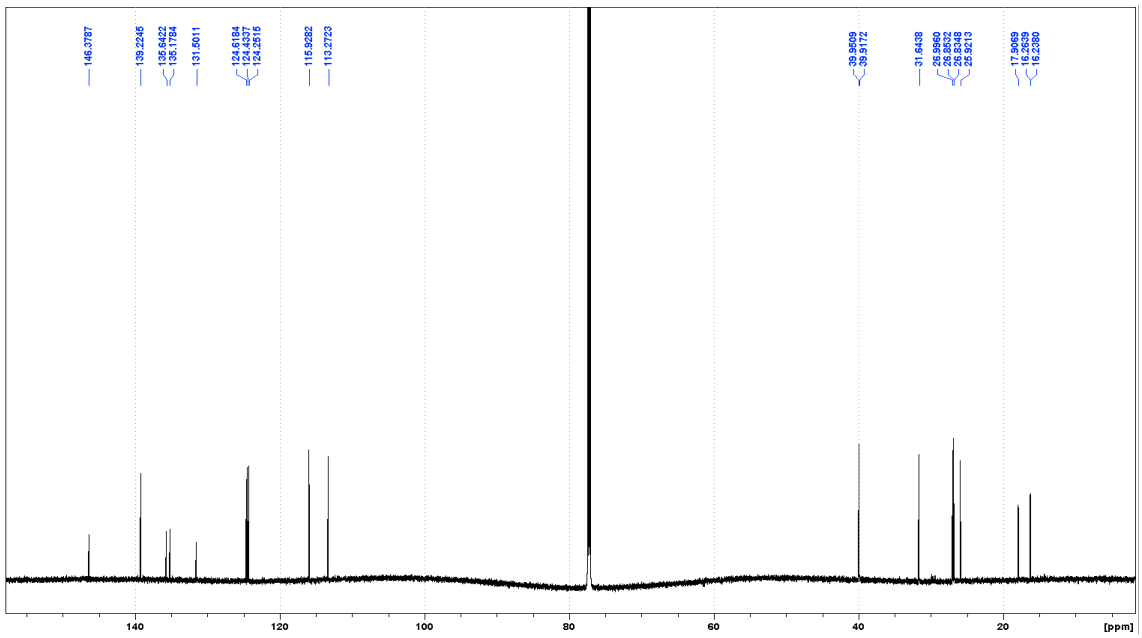


Fig. S7. (B) ^{13}C Spectrum of β-springene (**28**)

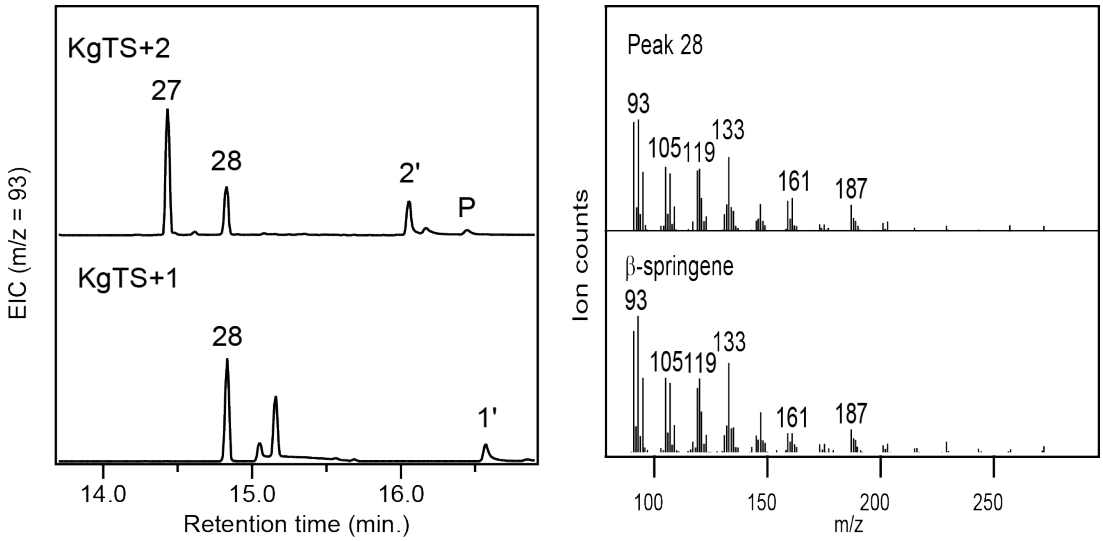


Fig. S8. β-Springene (**28**) afforded by KgTS upon co-expression with NNPPS was further verified by comparison to that produced by KgTS upon co-expression with the isoprenyl diphosphate synthase producing **1** (GGPPS).

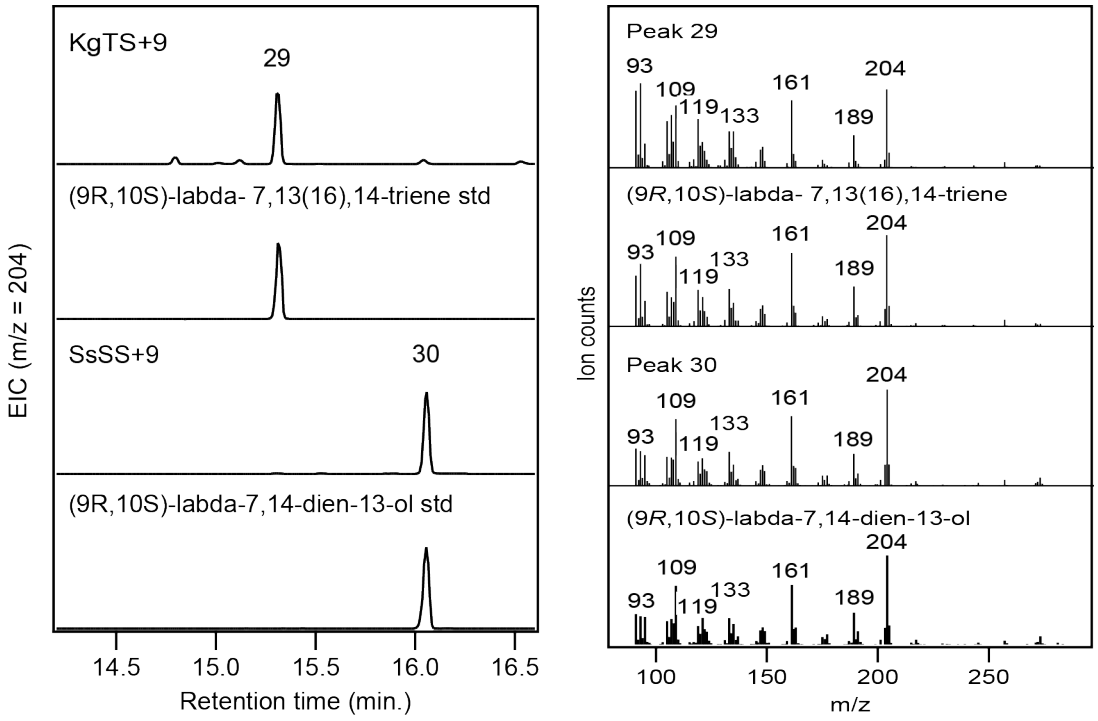
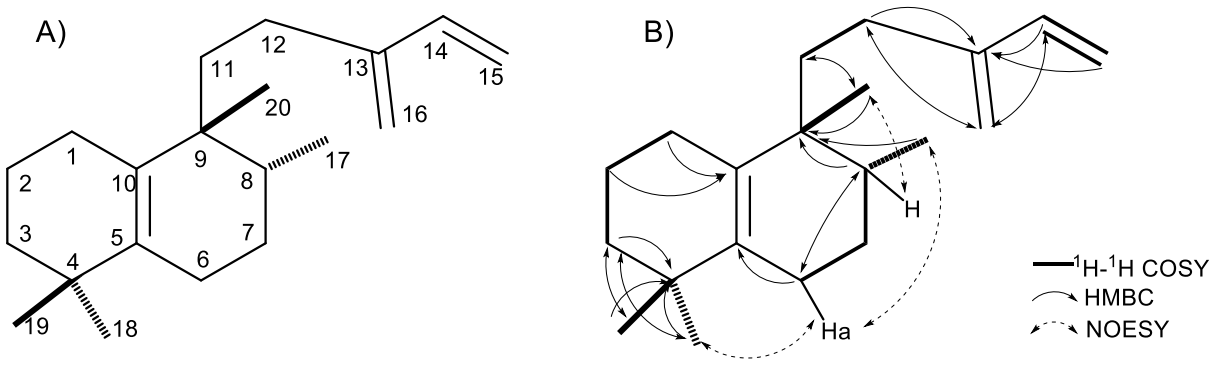


Fig. S9. Product identification by comparing both RT and MS from GC-MS analyses to authentic standards. Products **29** and **30** afforded by KgTS and SsSS reacting with *ent*-7-endo-CPP (**9**), respectively, were identified by comparison to the known *enantiomers* (Jia*, et al.*, 2016).



(8*R*,9*R*)-halima-5(10),13(16),14-triene (**31**)

Fig. S10. The product of KgTS upon co-expression with the DTC producing **18** (MvCPS1:W323F/F505Y). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-^{1}H COSY correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure.

Table S4. ^{1}H and ^{13}C NMR assignments for compound **31**, (8*R*,9*R*)-halima-5(10),13(16),14-triene (solvent $CDCl_3$).

Position	(8*R*,9*R*)-halima-5(10),13(16),14-triene (**31**)	
	δ_H	δ_C
1	1.91 (2H, m)	26.3
2	1.56 (2H, m)	20.2
3	1.40 (2H, m)	40.1
4		34.7
5		135.6
6a	2.03 (1H, m)	23.9
b	1.90 (1H, m)	
7a	1.58 (1H, m)	28.06
b	1.42 (1H, m)	
8	1.51 (1H, s)	38.1
9		40.2
10		132.6
11a	1.56 (1H, m)	35.1
b	1.40 (1H, m)	
12a	2.16 (1H, td, *J* = 13.5 Hz, 5.2 Hz)	28.08
b	2.03 (1H, m)	
13		148.3
14	6.33 (1H, dd, *J* = 17.7 Hz, 10.7	139.3
15a	5.19 (1H, d, *J* = 17.7 Hz)	113.2

b	5.02 (1H, d, *J* = 10.7 Hz)	
16	4.94 (2H, d, *J* = 4.2 Hz)	115.4
17	0.93 (3H, d, *J* = 6.8 Hz)	16.6
18	0.99 (3H, s)	28.6
19	0.97 (3H, s)	28.7
20	0.99 (3H, s)	27.0

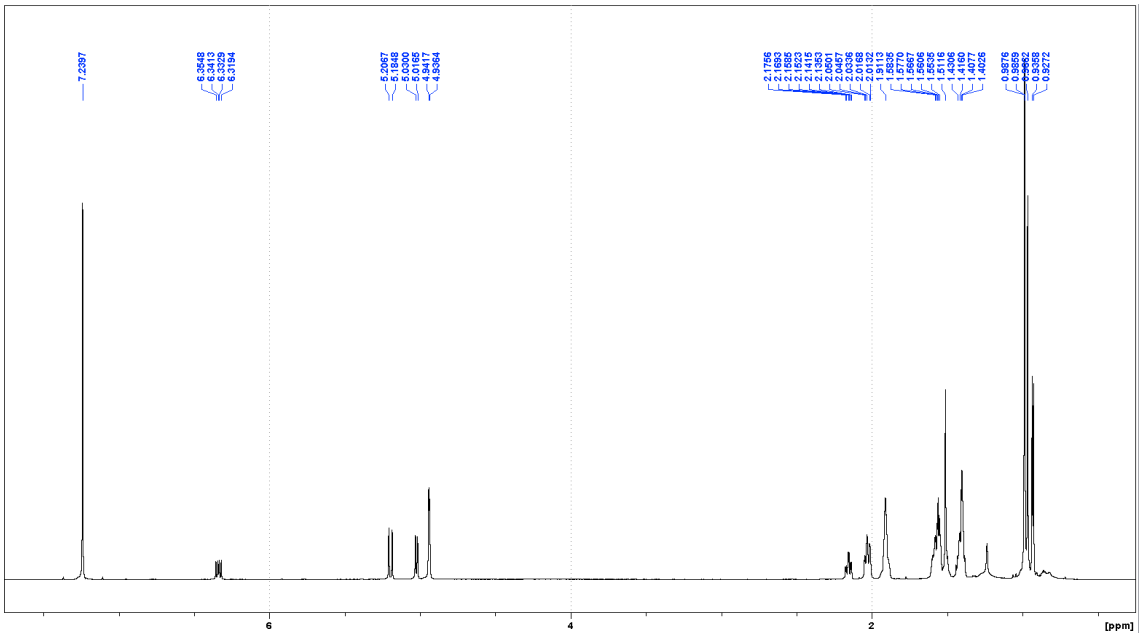


Fig. S11. (A) ^{1}H Spectrum of (8*R*,9*R*)-halima-5(10),13(16),14-triene (**31**)

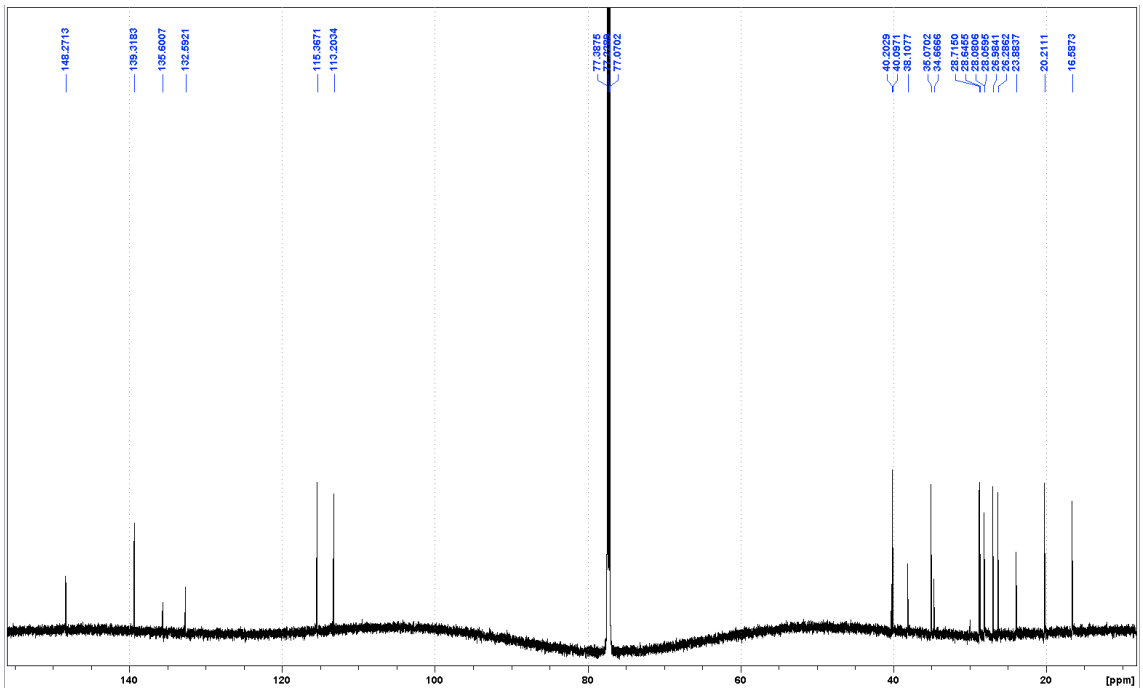


Fig. S11. (B) ^{13}C Spectrum of (8*R*,9*R*)-halima-5(10),13(16),14-triene (**31**)

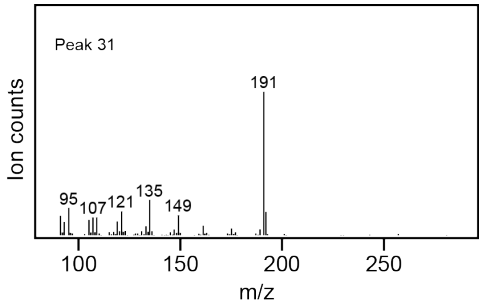
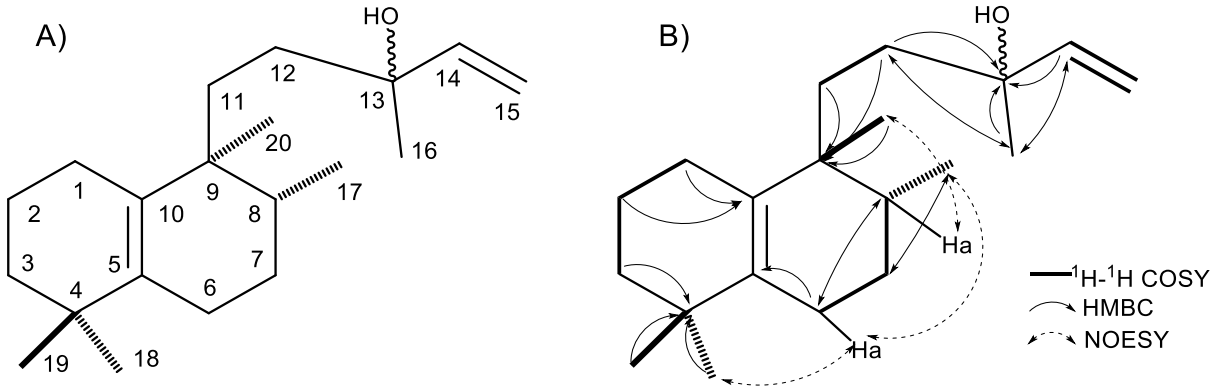


Fig. S12. MS of (8*R*,9*R*)-halima-5(10),13(16),14-triene (**31**)



(8*R*,9*R*)-halima-5(10),14-dien-13-ol (**32**)

Fig. S13. The product of SsSS upon co-expression with the DTC producing **18** (MvCPS1:W323F/F505Y). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-^{1}H COSY correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure. Splitting was observed for two protons and the majority of the carbons (Table S5 and Fig. S14), which is presumed to be due to a mixture of C13 stereoisomers (note that these are not separated in the non-chiral column used here).

Table S5. ^{1}H and ^{13}C NMR assignments for compound **32**, (8*R*,9*R*)-halima-5(10),14-dien-13-ol (solvent $CDCl_3$).

Position	(8*R*,9*R*)-halima-5(10),14-dien-13-ol (**32**)	
	δ_H	δ_C
1	1.88 (2H, m)	26.15/26.13
2	1.52 (2H, m)	20.21/20.20
3	1.39 (2H, m)	40.10
4		34.65
5		135.50/135.40
6a	2.00 (1H, m)	23.70/23.60
b	1.87 (1H, m)	
7a	1.55 (1H, m)	27.89/27.84
b	1.34 (1H, m)	
8	1.45 (1H, m)	37.90/37.81
9		39.71/39.64
10		132.60/132.55
11a	1.35 (1H, m)	29.47/29.40
b	1.24 (1H, m)	
12a	1.50 (1H, m)	38.45/38.39
b	1.36 (1H, m)	
13		73.87/73.74
14	5.86 (1H, dd, J = 17.5 Hz, 10.7 Hz)	145.40/145.38
15a	5.17 (1H, d, J = 17.5 Hz)	111.93/111.91
b	5.03 (1H, d, J = 10.7 Hz)	
16	1.24 (3H, s)	27.99/27.74
17	0.88 (3H, d, J = 6.8 Hz)	16.39/16.34
18	0.96 (3H, s)	28.79/28.78
19	0.95 (3H, s)	28.59/28.57
20	0.95 (3H, s)	27.18/27.15

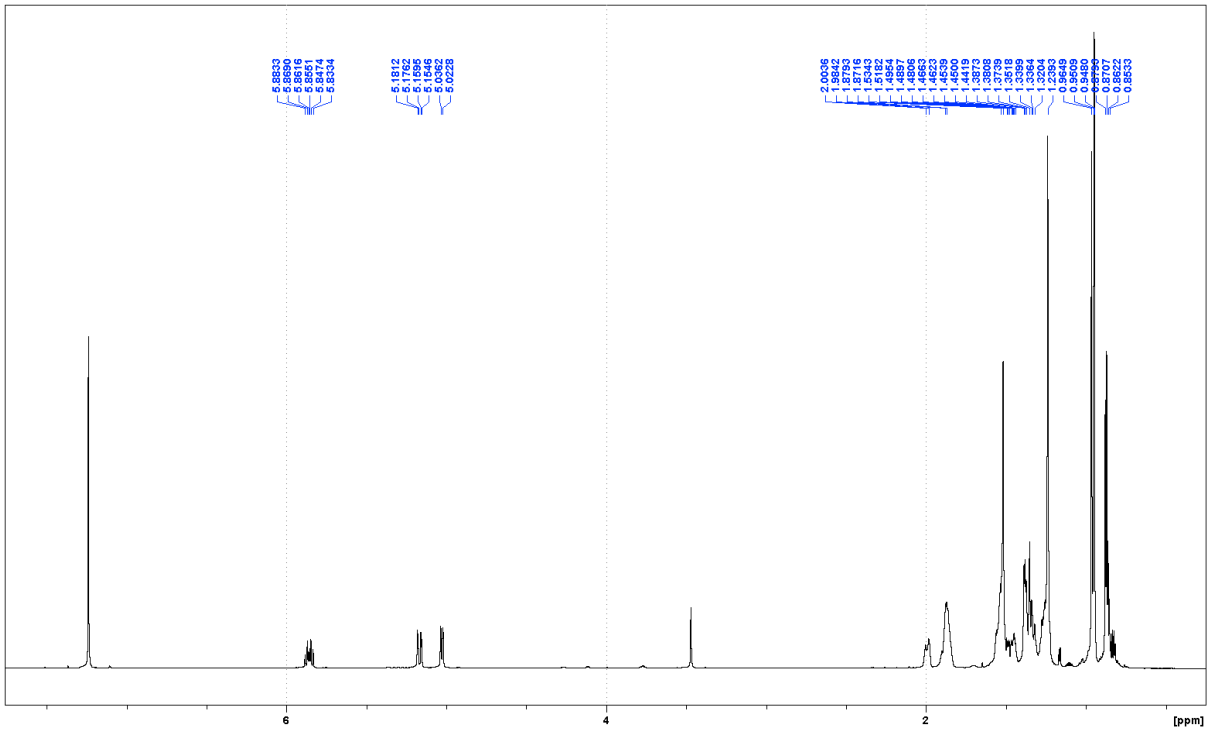


Fig. S14. (A) ^{1}H Spectrum of (8*R*,9*R*)-halima-5(10),14-dien-13-ol (**32**)

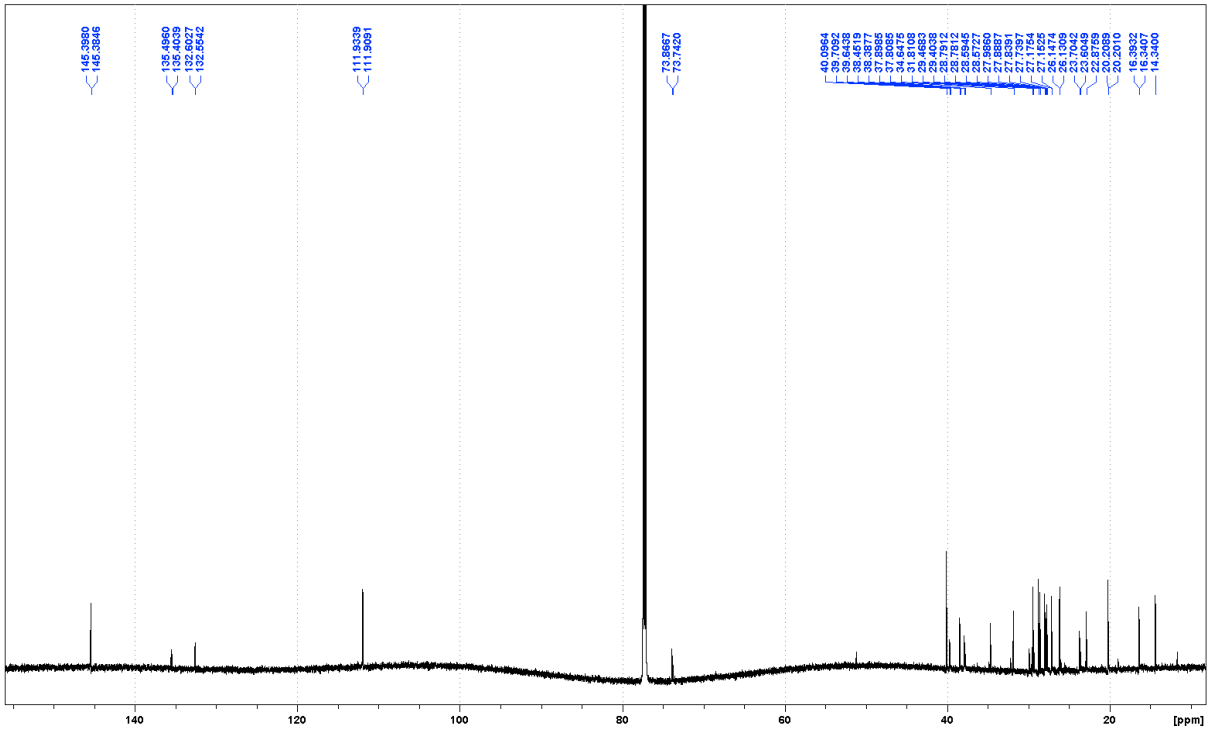


Fig. S14. (B) ^{13}C Spectrum of (8*R*,9*R*)-halima-5(10),14-dien-13-ol (**32**)

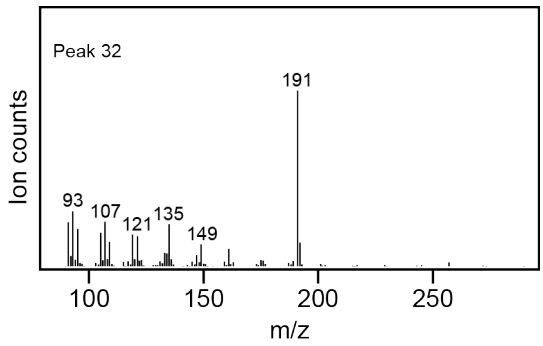
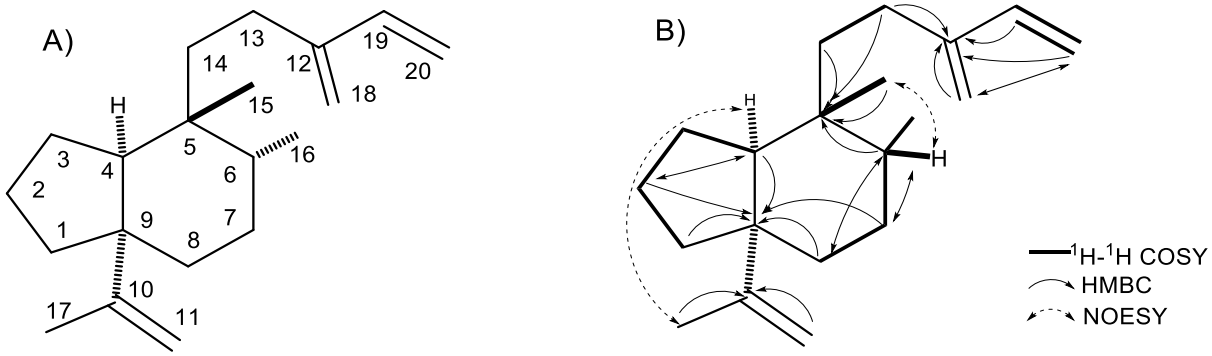


Fig. S15. MS of (8*R*,9R)-halima-5(10),14-dien-13-ol (**32**)



(4*S*,5*R*,6*R*,9*S*)-mutil-10,12(18),19-triene (**33**)

Fig. S16. The product of KgTS upon co-expression with the DTC producing **19** (CpPS:D649L). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-^{1}H COSY correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure.

Table S6. ^{1}H and ^{13}C NMR assignments for compound **33**, (4*S*,5*R*,6*R*,9*S*)-mutil-10,12(18),19-triene (solvent $CDCl_3$).

Position	(4*S*,5*R*,6*R*,9*S*)-mutil-10,12(18),19-triene (**33**)	
	δ_H	δ_C
1a	1.76 (1H, m)	41.2
b	1.20 (1H, m)	
2a	1.69 (1H, m)	21.9
b	1.54 (1H, m)	
3a	1.76 (1H, m)	27.4
b	1.68 (1H, m)	
4	2.18 (1H, t, J = 9.6 Hz)	48.2
5		37.6

6	1.68 (1H, m)	37.5
7a	1.50 (1H, m)	28.2
b	1.21 (1H, m)	
8a	1.82 (1H, d, *J* = 13.9 Hz)	31.6
b	1.43 (1H, m)	
9		49.4
10		151.4
11a	4.87 (1H, s)	109.6
b	4.76 (1H, s)	
12		147.4
13	2.10 (2H, m)	25.9
14a	1.68 (1H, m)	28.8
b	1.28 (1H, m)	
15	0.85 (3H, s)	25.5
16	0.79 (3H, d, *J* = 6.9 Hz)	16.1
17	1.72 (3H, s)	19.7
18	4.96 (2H, d, *J* = 18.3 Hz)	114.9
19	6.37 (1H, dd, *J* = 17.5 Hz, 10.8 Hz)	140.0
20 a	5.21 (1H, d, *J* = 17.5 Hz)	112.8
b	5.00 (1H, d, *J* = 10.8 Hz)	

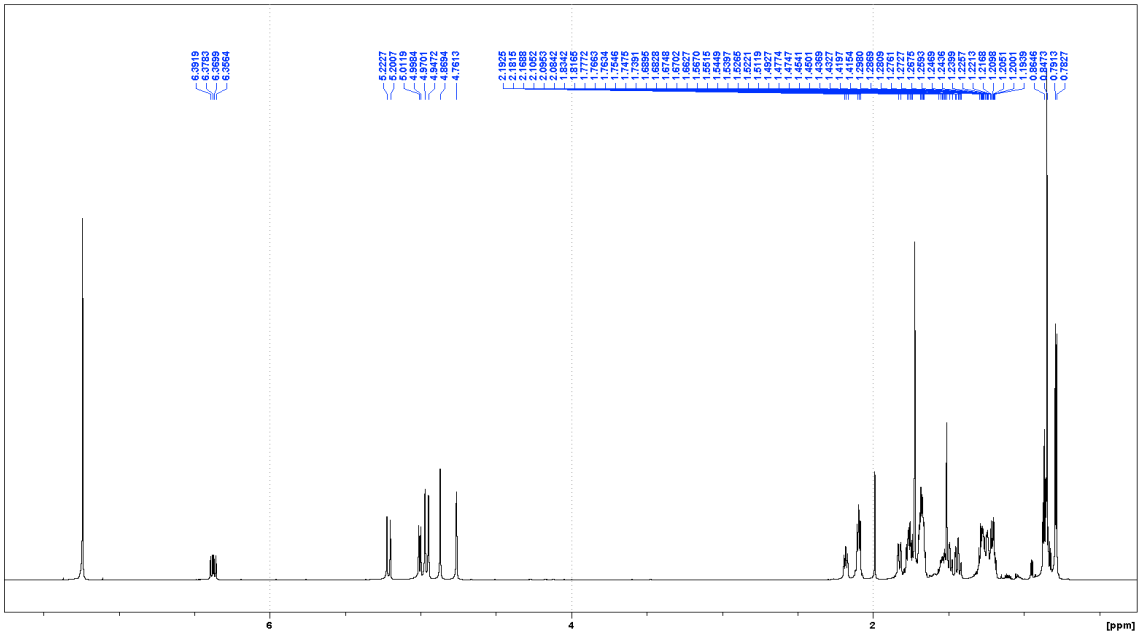


Fig. S17. (A) ^{1}H Spectrum of (4*S*,5*R*,6*R*,9*S*)-mutil-10,12(18),19-triene (**33**)

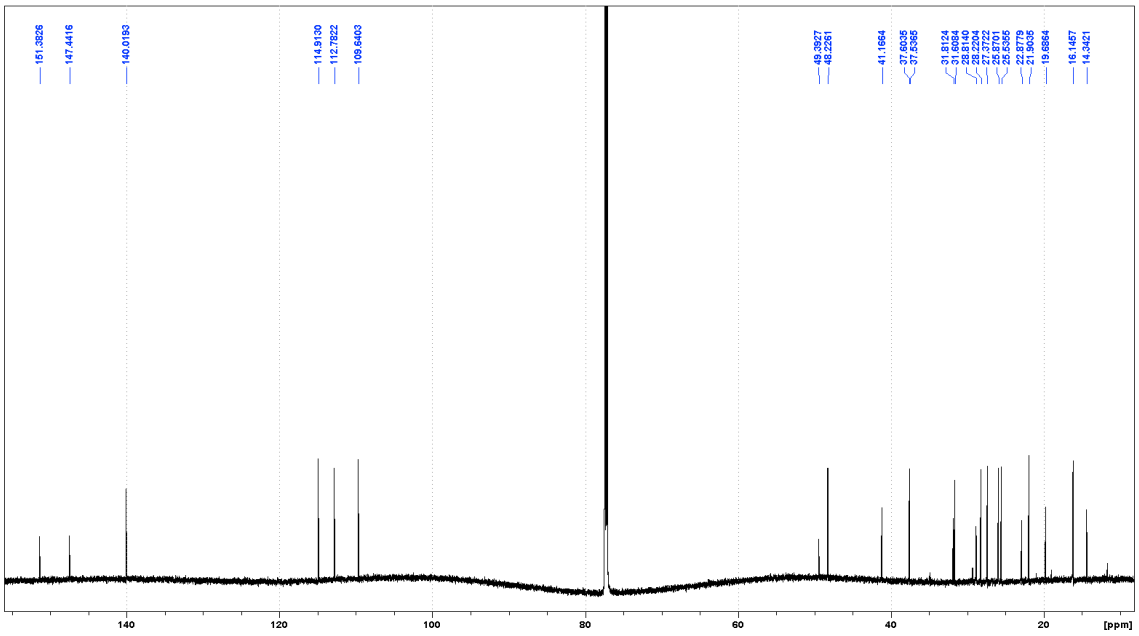


Fig. S17. (B) ^{13}C Spectrum of (4*S*,5*R*,6*R*,9*S*)-mutil-10,12(18),19-triene (**33**)

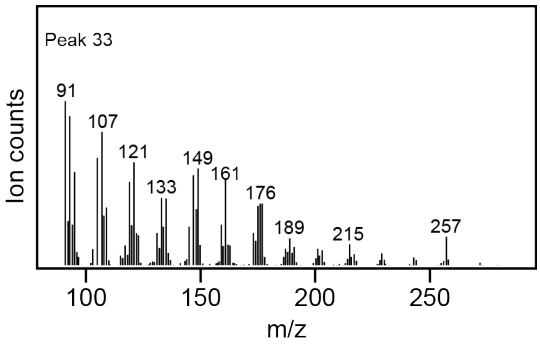


Fig. S18. MS of (4*S*,5*R*,6*R*,9*S*)-mutil-10,12(18),19-triene (**33**)

Table S7. Substrate conversion percentage values of eight DTSs with 19 substrates.

Substrate[a]	AbCAS[b]	ScLS[b]	SmMS[b]	AtKS[b]	OsKS[b]	SaPS[b]	BjKS[b]	EtKS[b]
1	-	58 ± 3	-	-	-	40 ± 5	27 ± 4	60 ± 6
2	-	10 ± 4	-	-	-	-	-	-
3	-	11 ± 3	-	-	-	-	13 ± 5	34 ± 5
4	-	-	-	-	-	-	-	-
5	26 ± 6	***100***	***100***	-	-	***100***	67 ± 8	85 ± 5
6	-	90 ± 4	-	***100***	***91± 4***	57 ± 8	***100***	***100***
7	-	80 ± 4	98 ± 2	-	-	100	80 ± 7	22 ± 4
8	-	100	-	-	-	38 ± 6	95 ± 4	95 ± 3
9	-	97 ± 2	-	20 ± 3	-	50 ± 4	94 ± 5	93 ± 4
10	***100***	100	98 ± 2	-	-	100	91 ± 6	100
11	-	100	-	100	91 ± 3	97 ± 2	100	100
12	-	100	-	-	-	50 ± 5	70 ± 3	72 ± 5
13	-	72 ± 6	-	-	-	-	-	-
14	-	77 ± 4	-	-	-	-	96 ± 4	87 ± 3
15	-	75 ± 6	-	-	-	-	-	-
16	-	100	-	-	-	-	-	-
17	-	98 ± 1	-	-	-	14 ± 4	15 ± 4	79 ± 4
18	-	93 ± 2	-	-	-	15 ± 3	25 ± 5	48 ± 6
19	-	17 ± 3	-	-	-	-	-	-

[a]Numbering here is consistent with that used throughout.

[b]Values are given as Mean ± SD from triplicate measurements. "-" means products were not detected. The conversion percentages of DTSs with their native substrates are highlighted in **bold** and *italicized* red text.

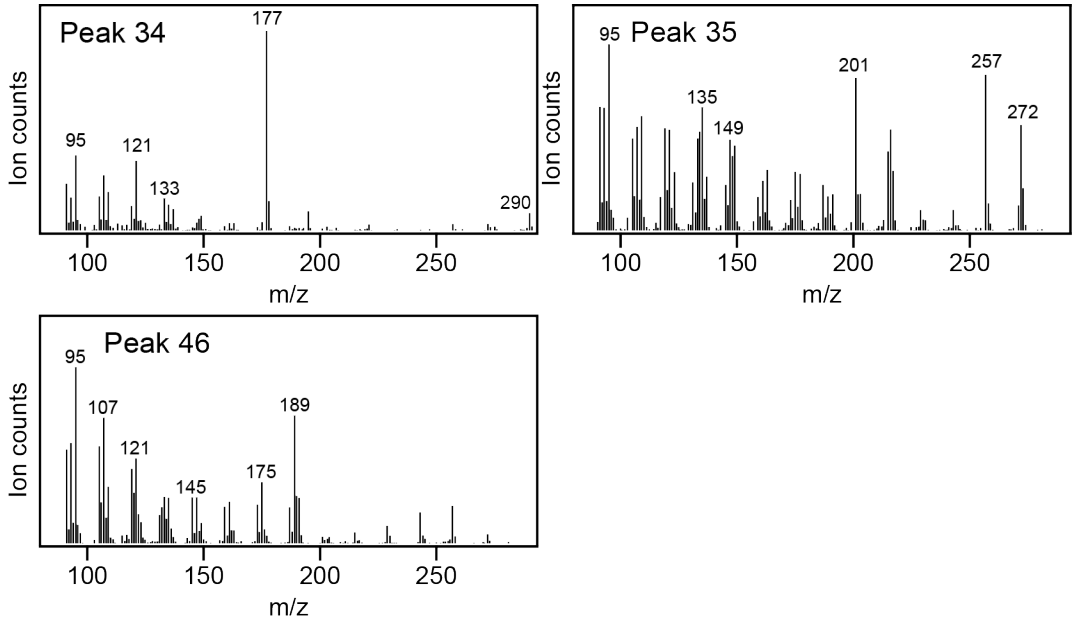


Fig. S19. Product identification by comparing both RT and MS from GC-MS analyses to authentic standards. For general reference the MS are presented here.

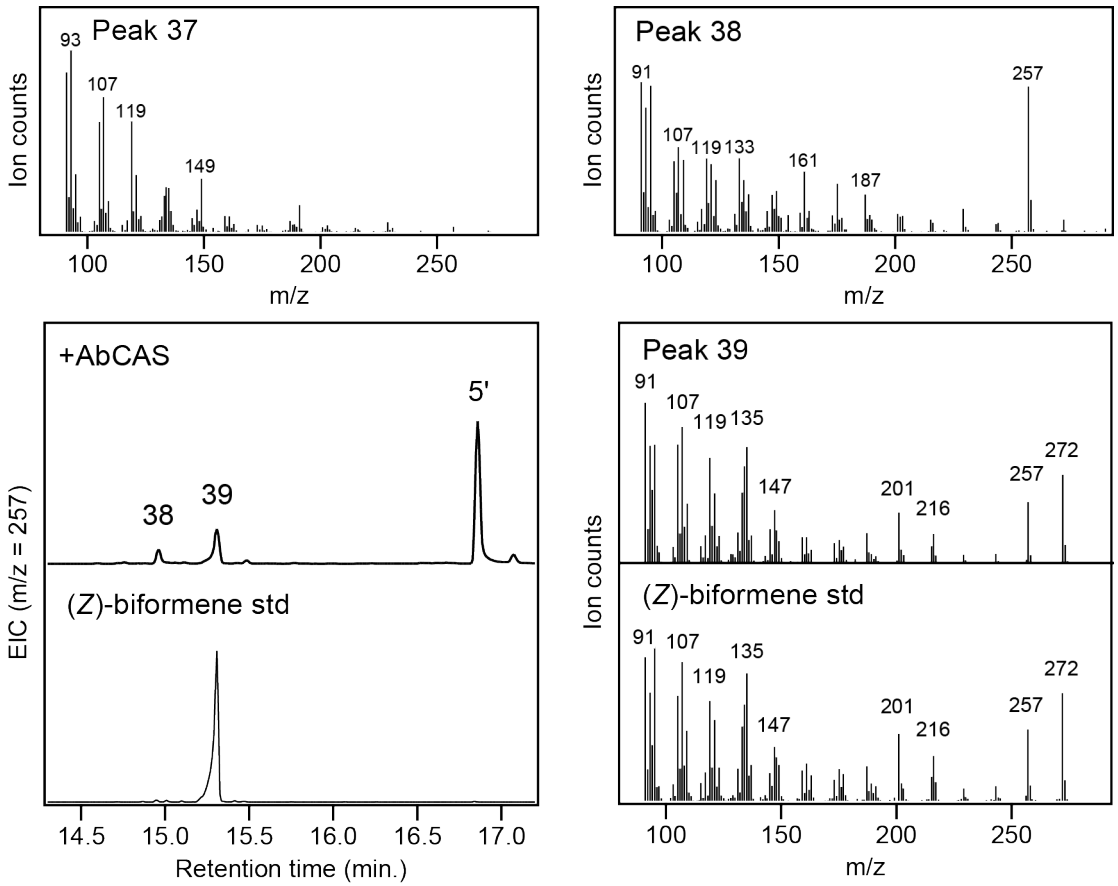


Fig. S20. Product identification by comparing both RT and MS from GC-MS analyses to authentic standards. New product **39** afforded by AbCAS reacting with CPP (**5**) was identified by such comparison to the known *enantiomer*. Otherwise for general reference only MS are shown here.

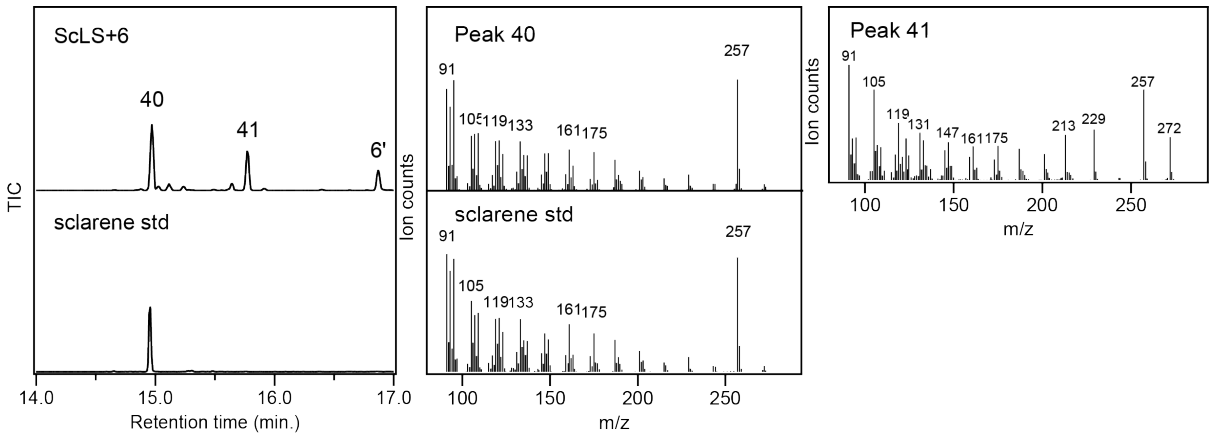


Fig. S21. Product identification by comparing both RT and MS from GC-MS analyses to authentic standards. New product **40** afforded by ScLS reacting with *ent*-CPP (**6**) was identified by such comparison to the known *enantiomer*. Otherwise for general reference only MS are shown here.

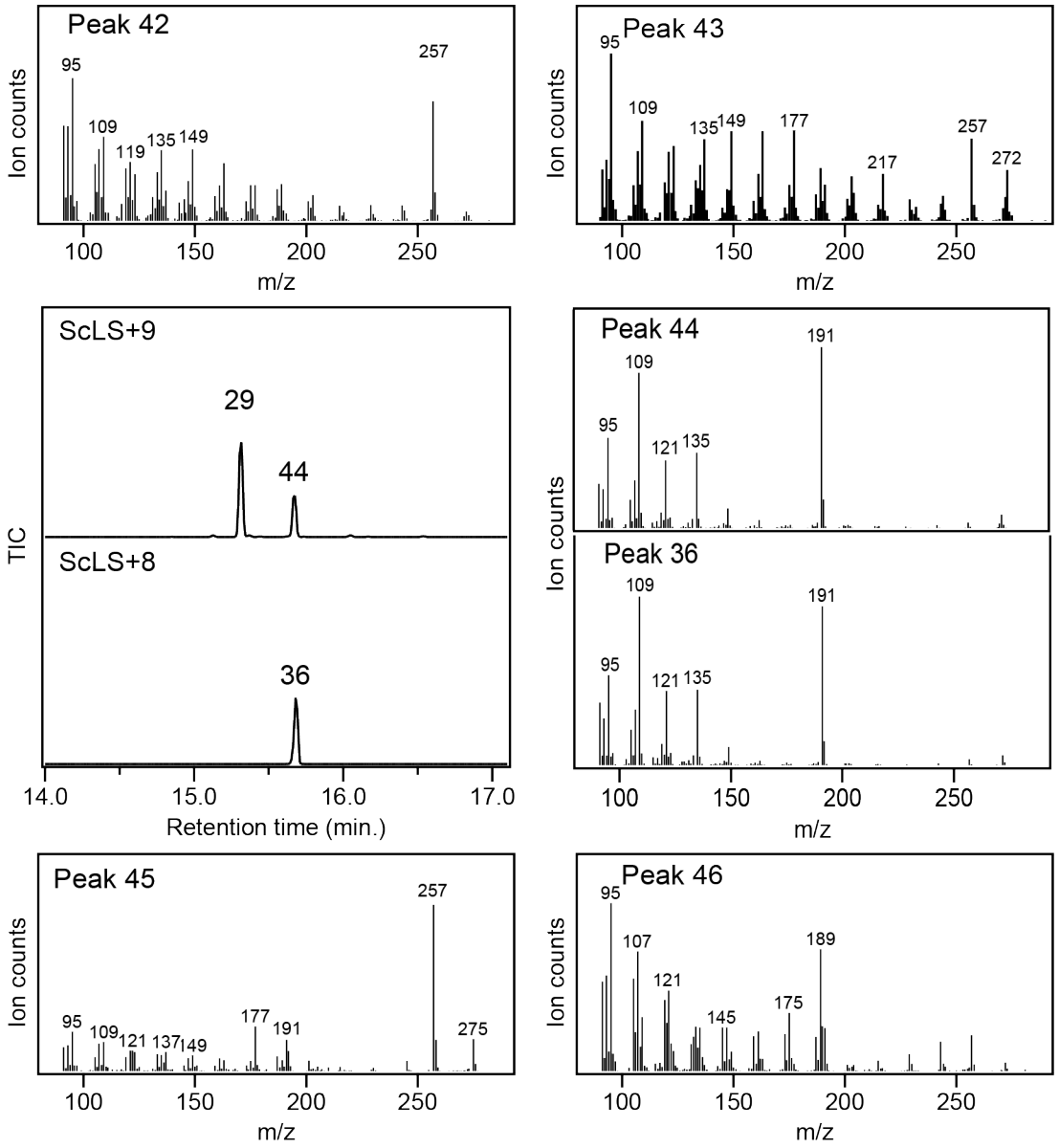


Fig. S22. Product identification by comparing both RT and MS from GC-MS analyses to authentic standards. New product **44** afforded by ScLS reacting with *ent*-7-endo-CPP (**9**) was identified by such comparison to the known *enantiomer*. Otherwise for general reference only MS are shown here.

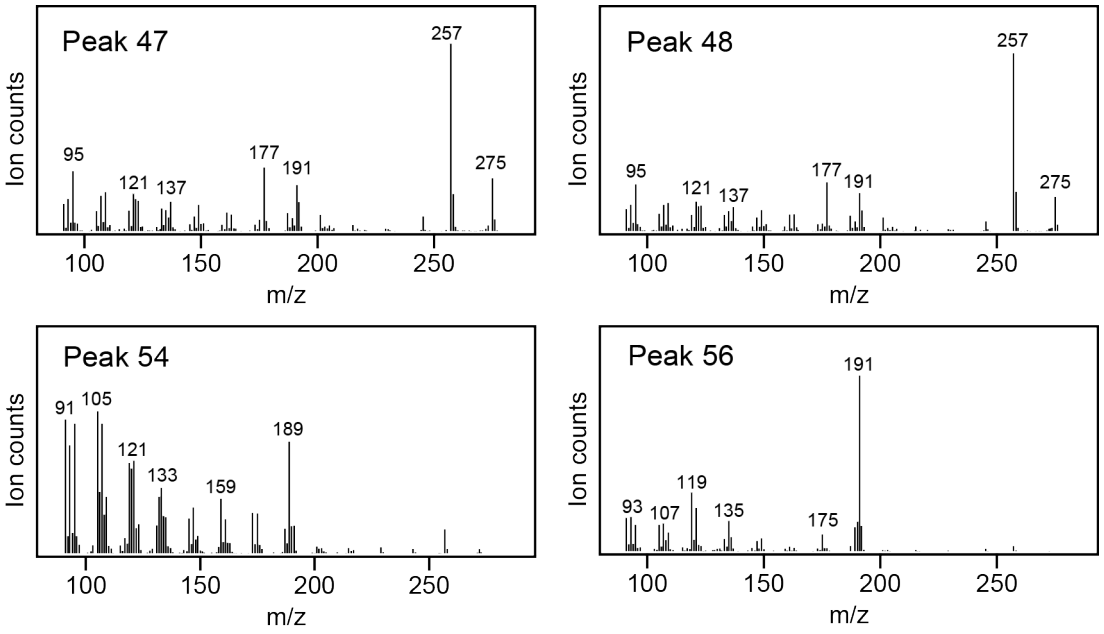
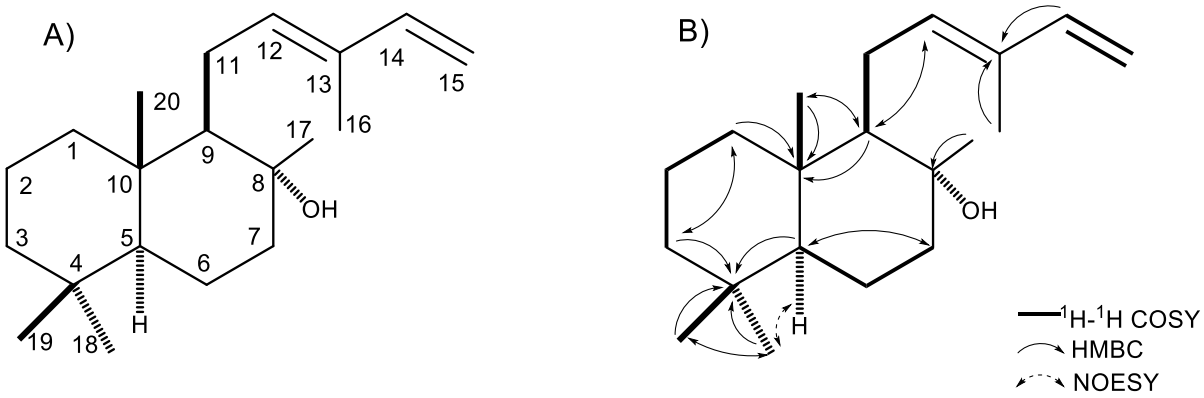


Fig. S23. Product identification by comparing both RT and MS from GC-MS analyses to authentic standards. For general reference the MS are presented here.



(8R,9R,10S)-labda-12*E*,14-dien-8-ol (**49**)

Fig. S24. The product of ScLS upon co-expression with the DTC producing **10** (NgCLS). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-^{1}H COSY correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure.

Table S8. ^{1}H and ^{13}C NMR assignments for compound **49**, (8*R*,9*R*,10*S*)-labda-12*E*,14-dien-8-ol (solvent $CDCl_3$).

Position	(8*R*,9*R*,10*S*)-labda-12*E*,14-dien-8-ol (**49**)	
	δ_H	δ_C
1a	1.62 (1H, m)	40.3
b	0.90 (1H, m)	
2a	1.56 (1H, m)	18.8
b	1.40 (1H, m)	
3a	1.35 (1H, d, *J* = 13.1 Hz)	42.1
b	1.12 (1H, td, *J* = 13.5 Hz, 4.3 Hz)	
4		33.5
5	0.92 (1H, m)	56.3
6a	1.63 (1H, m)	20.6
b	1.24 (1H, m)	
7a	1.84 (1H, dt, *J* = 12.6 Hz, 3.4 Hz)	44.3
b	1.41 (1H, m)	
8		74.3
9	1.31 (1H, t, *J* = 5.4 Hz)	62.4
10		39.2
11a	2.33 (1H, m)	24.3
b	2.17 (1H, m)	
12	5.58 (1H, t, *J* = 7.3 Hz)	136.3
13		132.8
14	6.33 (1H, dd, *J* = 17.4 Hz, 10.6 Hz)	141.8
15a	5.04 (1H, d, *J* = 17.4 Hz)	110.6
b	4.89 (1H, d, *J* = 10.6 Hz)	
16	1.77 (3H, s)	12.1
17	1.17 (3H, s)	24.6
18	0.85 (3H, s)	33.7
19	0.78 (3H, s)	21.8
20	0.82 (3H, s)	15.7

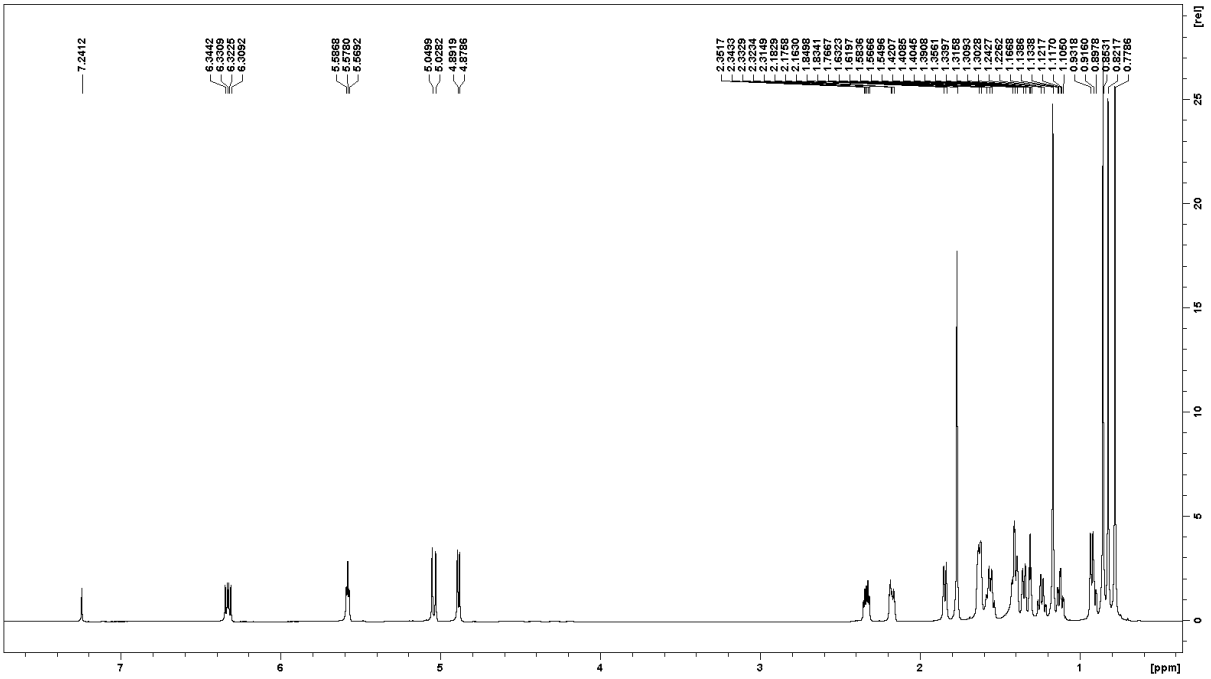


Fig. S25. (A) ^{1}H Spectrum of (8*R*,9*R*,10*S*)-labda-12*E*,14-dien-8-ol (**49**)

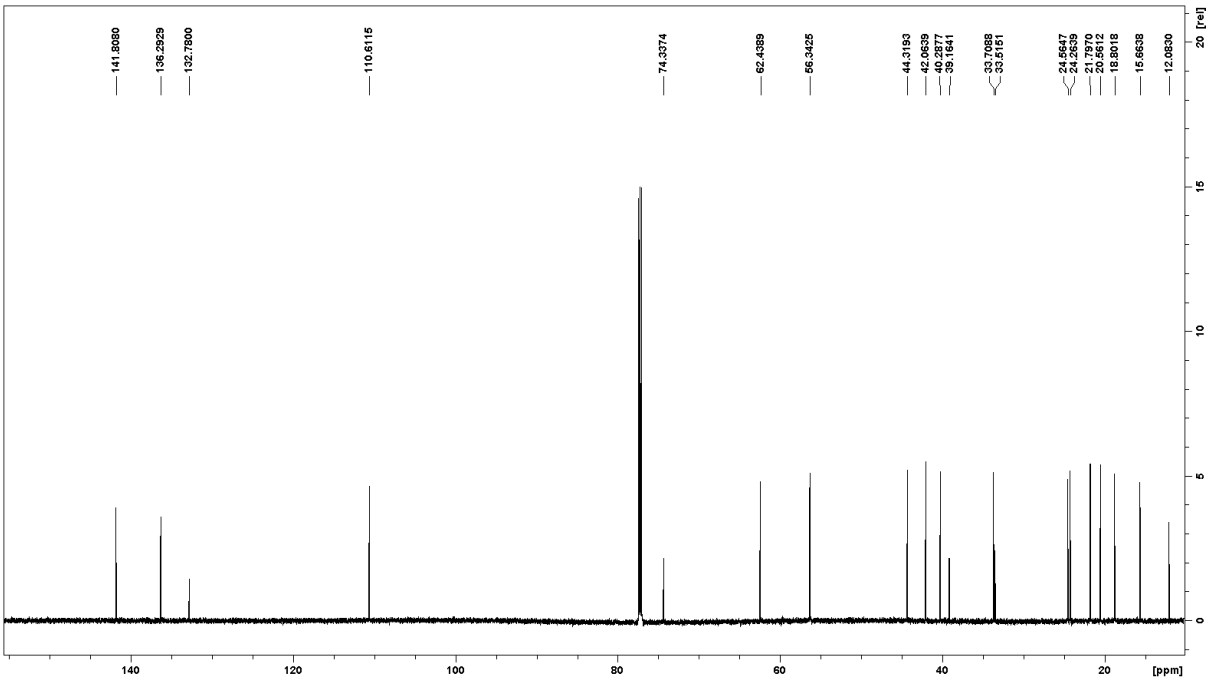


Fig. S25. (B) ^{13}C Spectrum of (8*R*,9*R*,10*S*)-labda-12*E*,14-dien-8-ol (**49**)

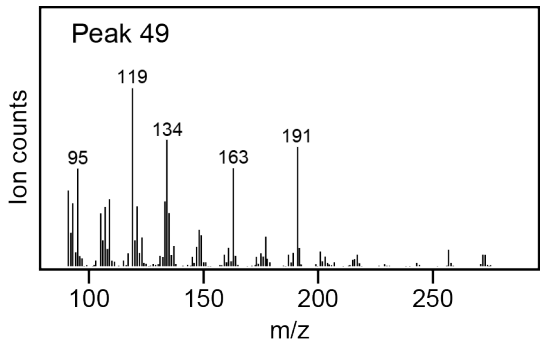
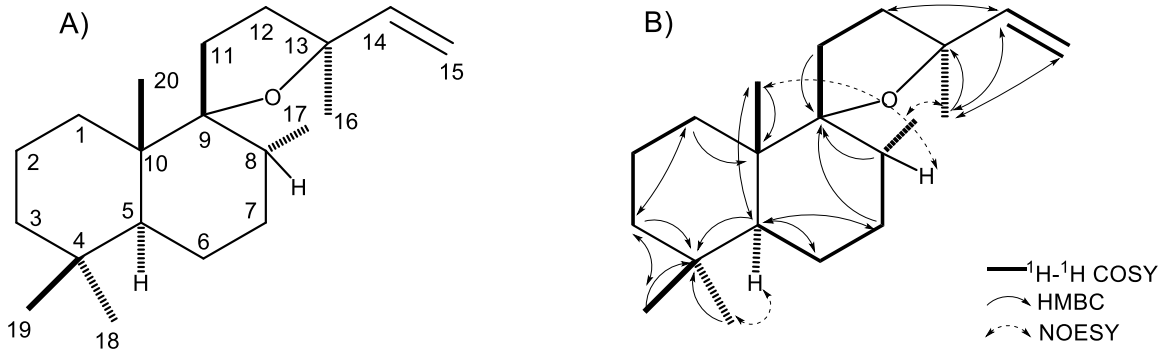


Fig. S26. MS of (8*R*,9*R*,10*S*)-labda-12*E*,14-dien-8-ol (**49**)



(8*R*,9*R*,10*S*,13*S*)-labda-9,13-epoxy-14-ene (**50**)

Fig. S27. The major product of ScLS upon co-expression with the DTC producing **12** (MvCPS1). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-^{1}H COSY correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure.

Table S9. ^{1}H and ^{13}C NMR assignments for compound **50**, (8*R*,9*R*,10*S*,13*S*)-labda-9,13-epoxy-14-ene (solvent $CDCl_3$)

Position	(8*R*,9*R*,10*S*,13*S*)-labda-9,13-epoxy-14-ene (**50**)	
	δ_H	δ_C
1a	1.56 (1H, m)	33.2
b	1.41 (1H, m)	
2a	1.56 (1H, m)	19.3
b	1.46 (1H, m)	
3a	1.30 (1H, m)	42.3
b	1.16 (1H, m)	
4		33.8
5	1.44 (1H, m)	47.3
6a	1.53 (1H, m)	22.26
b	1.26 (1H, m)	
7a	1.40 (1H, m)	32.0
b	1.32 (1H, m)	
8	1.67 (1H, m)	37.1
9		93.2
10		42.4
11a	2.09 (1H, m)	30.0
b	1.58 (1H, m)	
12a	1.96 (1H, m)	37.7
b	1.64 (1H, m)	
13		84.0
14	6.05 (1H, dd, J = 17.4 Hz, 10.8 Hz)	146.2
15a	5.14 (1H, d, J = 17.4 Hz)	110.0
b	4.93 (1H, d, J = 10.8 Hz)	
16	1.28 (3H, s)	29.2
17	0.78 (3H, d, J = 6.9 Hz)	18.4
18	0.84 (3H, s)	33.5
19	0.78 (3H, s)	22.33
20	0.86 (3H, s)	17.8

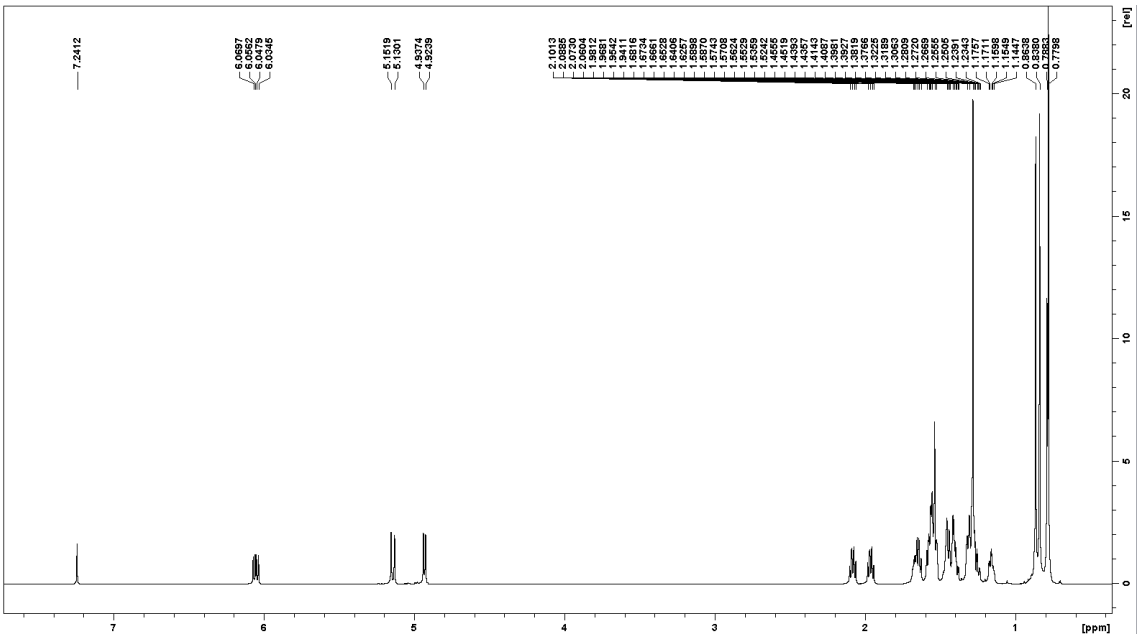


Fig. S28. (A) ^{1}H Spectrum of (8*R*,9*R*,10*S*,13*S*)-labda-9,13-epoxy-14-ene (**50**)

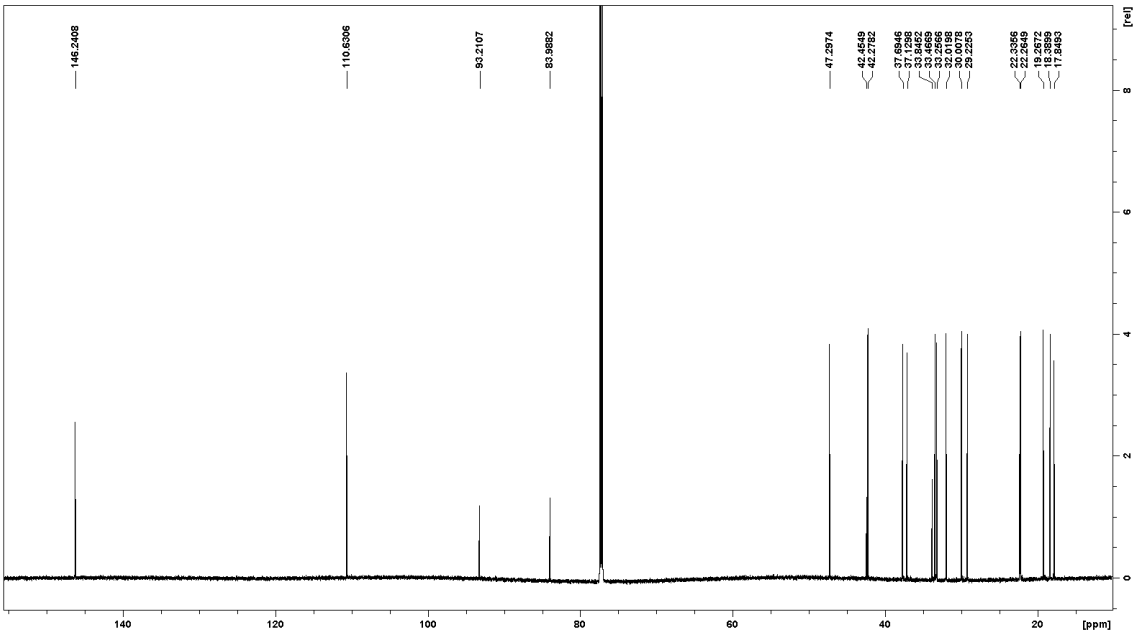


Fig. S28. (B) ^{13}C Spectrum of (8*R*,9*R*,10*S*,13*S*)-labda-9,13-epoxy-14-ene (**50**)

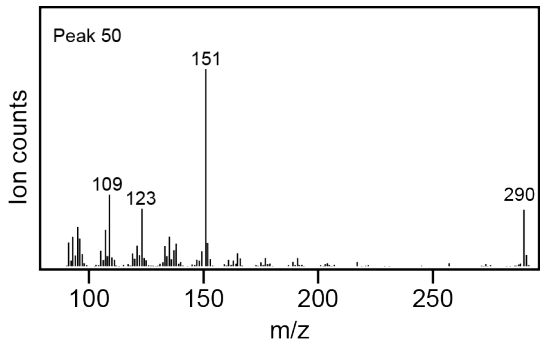
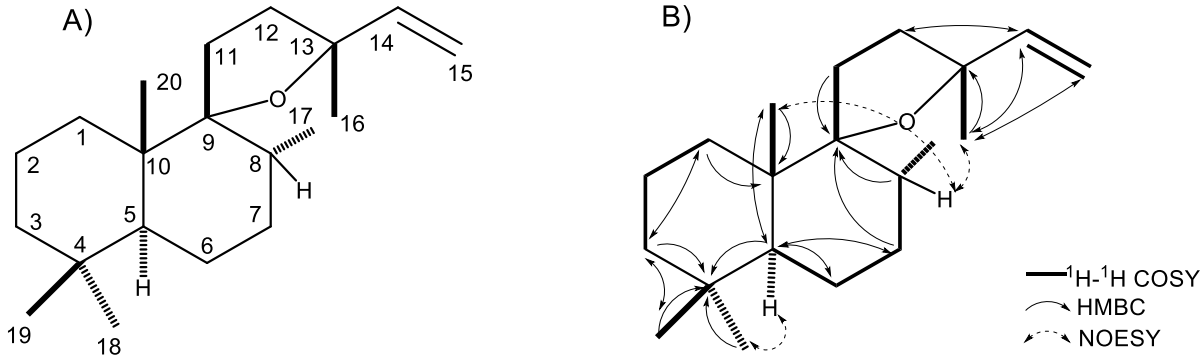


Fig. S29. MS of (8*R*,9*R*,10*S*,13*S*)-labda-9,13-epoxy-14-ene (**50**)



(8*R*,9*R*,10*S*,13*R*)-labda-9,13-epoxy-14-ene (**51**)

Fig. S30. The minor product of ScLS upon co-expression with the DTC producing **12** (MvCPS1). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-^{1}H COSY correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure.

Table S10. ^{1}H and ^{13}C NMR assignments for compound **51**, (8*R*,9*R*,10*S*,13*R*)-labda-9,13-epoxy-14-ene (solvent $CDCl_3$)

Position	(8*R*,9*R*,10*S*,13*R*)-labda-9,13-epoxy-14-ene (**51**)	
	δ_H	δ_C
1a	1.47 (1H, m)	33.8
b	1.17 (1H, m)	
2a	1.48 (1H, m)	19.1
b	1.40 (1H, m)	
3a	1.27 (1H, m)	42.2
b	1.12 (1H, m)	
4		33.6
5	1.49 (1H, m)	46.6
6a	1.51 (1H, m)	22.1
b	1.26 (1H, m)	
7	1.37 (2H, m)	32.0
8	1.68 (1H, m)	35.8
9		93.3
10		43.0
11a	1.94 (1H, m)	29.7
b	1.69 (1H, m)	
12a	1.95 (1H, m)	36.7
b	1.72 (1H, m)	
13		83.9
14	6.05 (1H, dd, J = 17.5 Hz, 10.8 Hz)	146.4
15a	5.05 (1H, d, J = 17.5 Hz)	110.0
b	4.91 (1H, d, J = 10.8 Hz)	
16	1.30 (3H, s)	29.1
17	0.87 (3H, d, J = 6.6 Hz)	18.4
18	0.84 (3H, s)	33.5
19	0.77 (3H, s)	22.2
20	0.84 (3H, s)	17.9

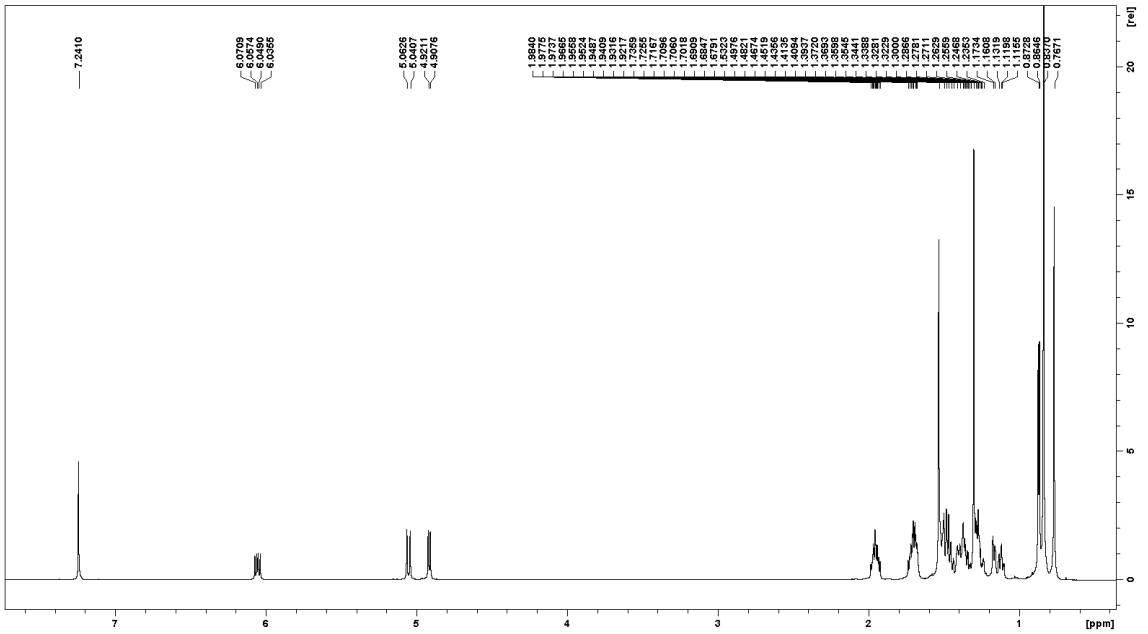


Fig. S31. (A) ^{1}H Spectrum of (8*R*,9*R*,10*S*,13*R*)-labda-9,13-epoxy-14-ene (**51**)

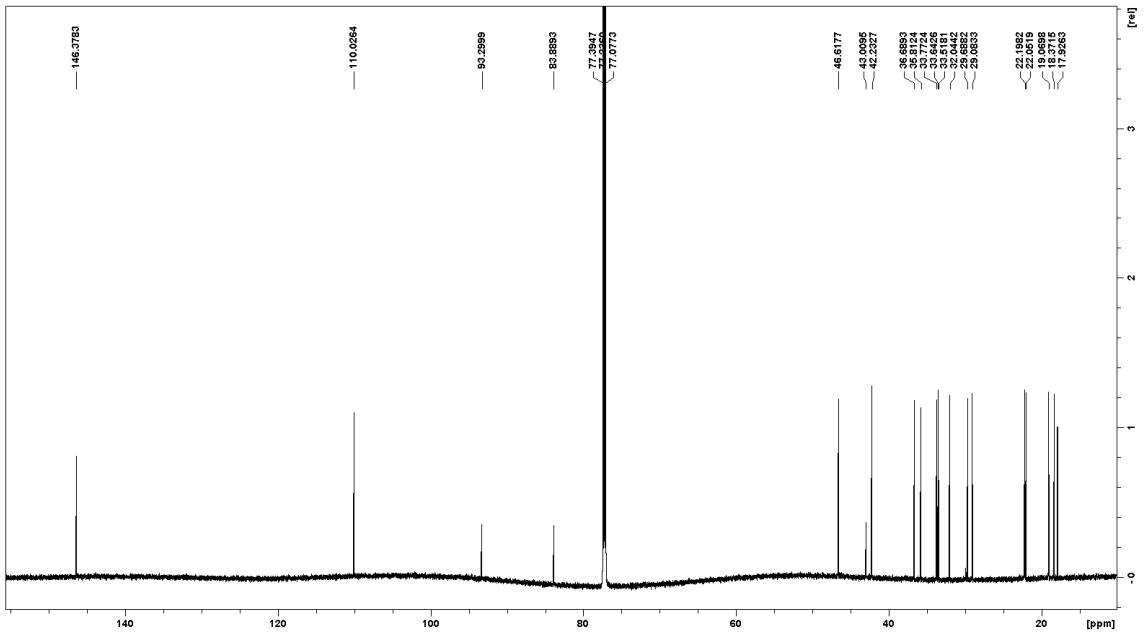


Fig. S31. (B) ^{13}C Spectrum of (8*R*,9*R*,10*S*,13*R*)-labda-9,13-epoxy-14-ene (**51**)

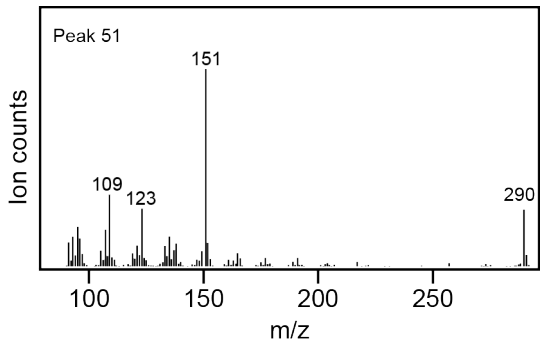
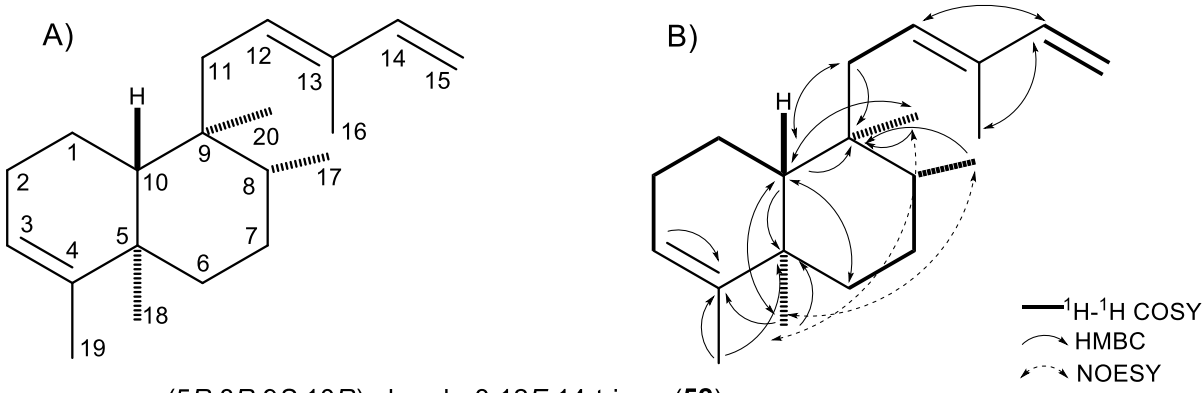


Fig. S32. MS of (8*R*,9*R*,10*S*,13*R*)-labda-9,13-epoxy-14-ene (**51**)



(5*R*,8*R*,9*S*,10*R*)-cleroda-3,12*E*,14-triene (**52**)

Fig. S33. The product of ScLS upon co-expression with the DTC producing **14** (AtCPS:H263Y). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-^{1}H COSY correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure.

Table S11. ^{1}H and ^{13}C NMR assignments for compound **52**, (5*R*,8*R*,9*S*,10*R*)-cleroda-3,12*E*,14-triene (solvent $CDCl_3$)

Position	(5*R*,8*R*,9*S*,10*R*)-cleroda-3,12*E*,14-triene (**52**)	
	δ_H	δ_C
1a	1.58 (1H, m)	19.0
b	1.42 (1H, m)	
2a	1.99 (1H, m)	27.1
b	1.90 (1H, m)	
3	5.13 (1H, brs)	120.8
4		144.6

5		38.5
6a	1.65 (1H, m)	36.7
b	1.13 (1H, m)	
7	1.37 (2H, m)	27.8
8	1.41 (1H, m)	37.1
9		40.7
10	1.32 (1H, d, *J* = 12.0 Hz)	47.5
11a	2.20 (1H, dd, J = 15.6 Hz, 7.9 Hz)	37.0
b	2.10 (1H, dd, J = 15.6 Hz, 7.7 Hz)	
12	5.47 (1H, t, *J* = 7.5 Hz)	130.0
13		135.0
14	6.36 (1H, dd, *J* = 17.4 Hz, 10.6 Hz)	142.3
15a	5.04 (1H, d, *J* = 17.4 Hz)	110.4
b	4.89 (1H, d, *J* = 10.6 Hz)	
16	1.73 (3H, s)	12.1
17	0.82 (3H, d, J = 6.4 Hz)	16.6
18	0.97 (3H, s)	20.0
19	1.55 (3H, s)	18.24
20	0.74 (3H, s)	18.15

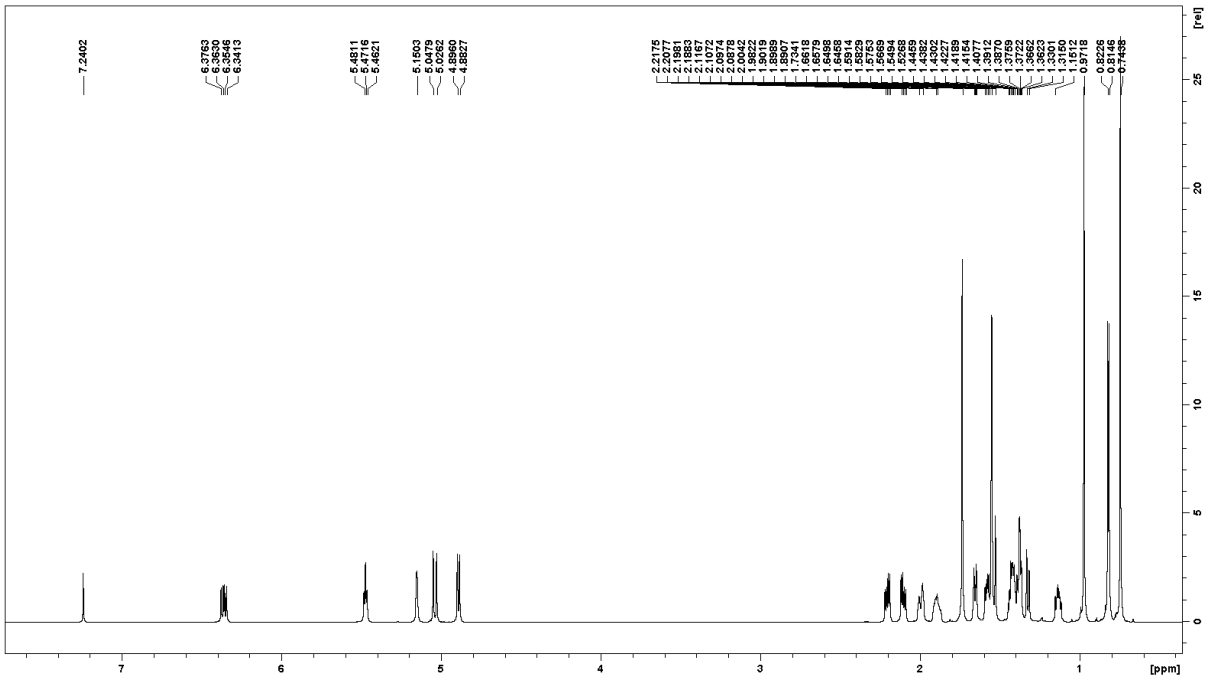


Fig. S34. (A) ^{1}H Spectrum of (5*R*,8*R*,9*S*,10*R*)-cleroda-3,12*E*,14-triene (**52**)

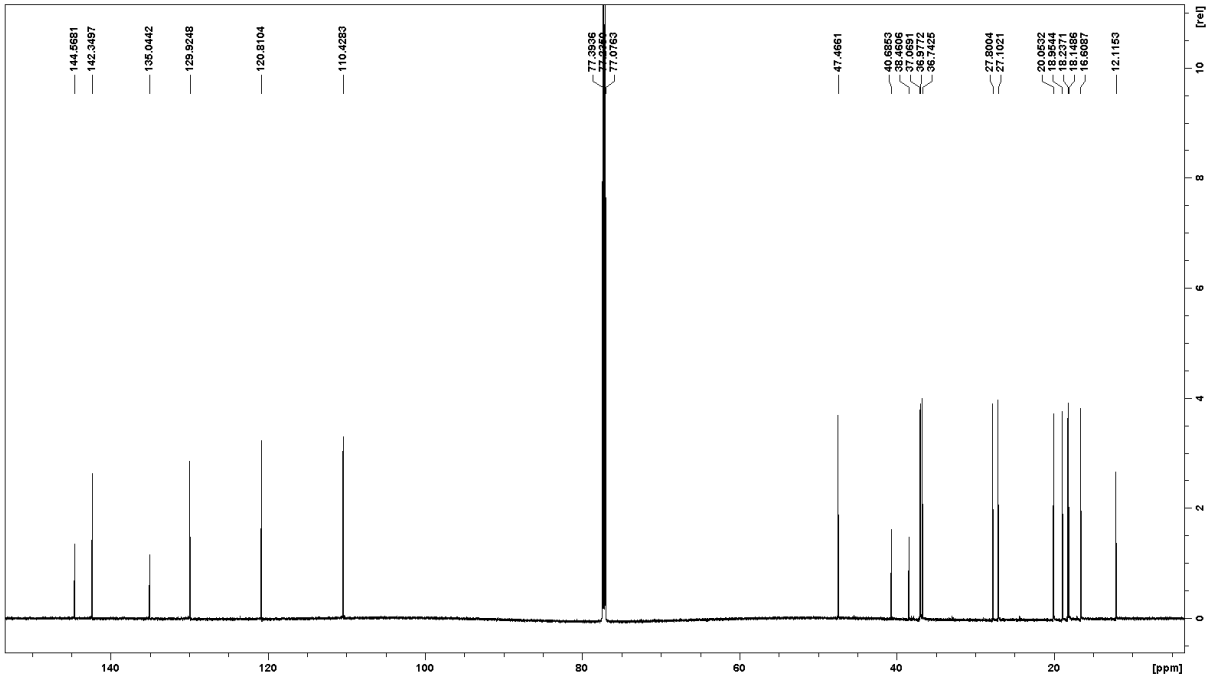


Fig. S34. (B) ^{13}C Spectrum of (5*R*,8*R*,9*S*,10*R*)-cleroda-3,12*E*,14-triene (**52**)

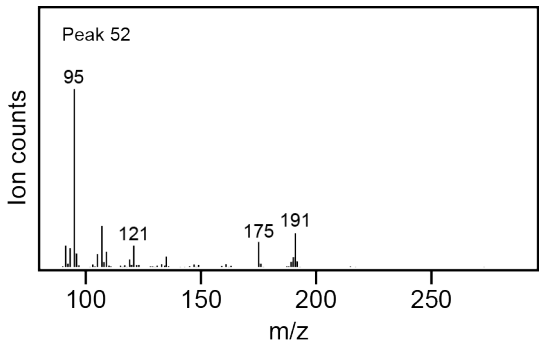
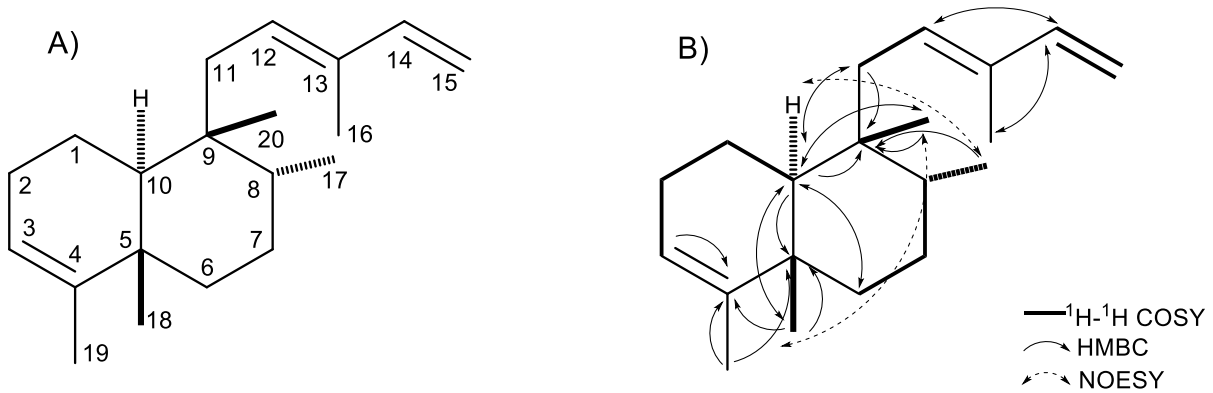


Fig. S35. MS of (5*R*,8*R*,9*S*,10*R*)-cleroda-3,12*E*,14-triene (**52**)



(5*S*,8*R*,9*R*,10*S*)-cleroda-3,12*E*,14-triene (**53**)

Fig. S36. The product of ScLS upon co-expression with the DTC producing **15** (KgTPS). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-^{1}H COSY

correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure.

Table S12. ^{1}H and ^{13}C NMR assignments for compound **53**, (5*S*,8*R*,9*R*,10*S*)-cleroda-3,12*E*,14-triene (solvent $CDCl_3$)

Position	(5*S*,8*R*,9*R*,10*S*)-cleroda-3,12*E*,14-triene (**53**)	
	δ_H	δ_C
1a	1.63 (1H, m)	18.28
b	1.43 (1H, m)	
2a	2.05 (1H, m)	27.1
b	1.99 (1H, m)	
3	5.15 (1H, brs)	120.4
4		144.8
5		38.6
6a	1.92 (1H, m)	25.8
b	1.25 (1H, m)	
7	1.42 (2H, m)	30.5
8	1.60 (1H, m)	35.7
9		38.9
10	1.49 (1H, m)	45.2
11a	2.29 (1H, dd, J = 15.4 Hz, 7.6 Hz)	38.2
b	1.82 (1H, dd, J = 15.4 Hz, 6.5 Hz)	
12	5.59 (1H, t, *J* = 7.1 Hz)	130.5
13		135.4
14	6.38 (1H, dd, *J* = 17.4 Hz, 10.7 Hz)	142.2
15a	5.04 (1H, d, *J* = 17.4 Hz)	110.2
b	4.89 (1H, d, *J* = 10.7 Hz)	
16	1.71 (3H, s)	12.3
17	0.91 (3H, d, J = 7.3 Hz)	15.2
18	1.02 (3H, s)	20.6
19	1.57 (3H, s)	18.27
20	0.90 (3H, s)	21.2

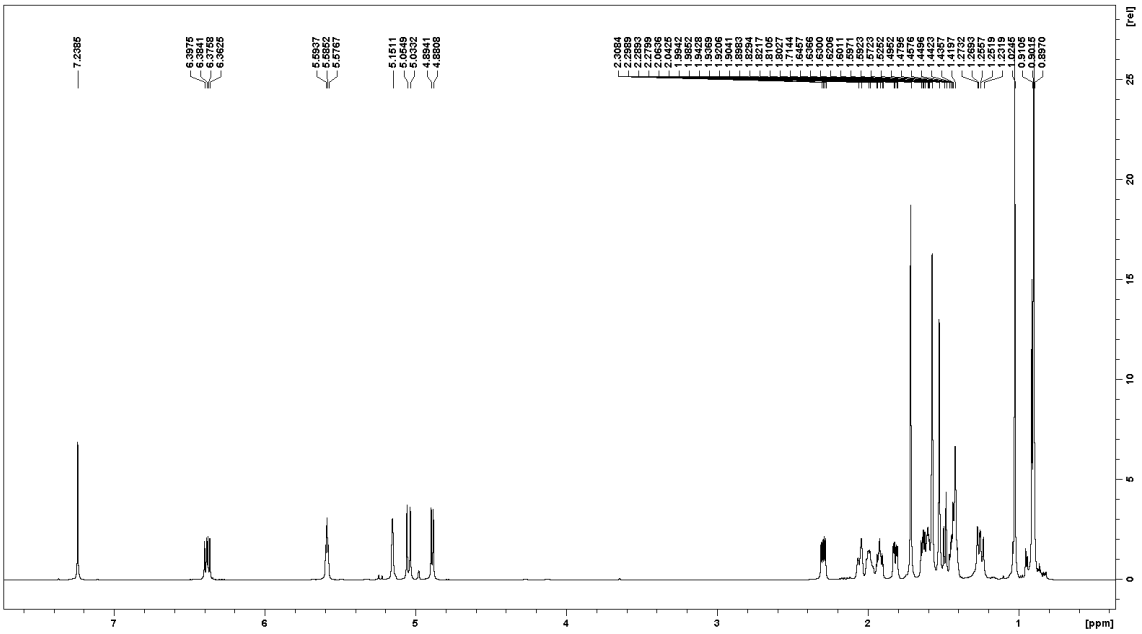


Fig. S37. (A) ^{1}H Spectrum of (5*S*,8*R*,9*R*,10*S*)-cleroda-3,12*E*,14-triene (**53**)

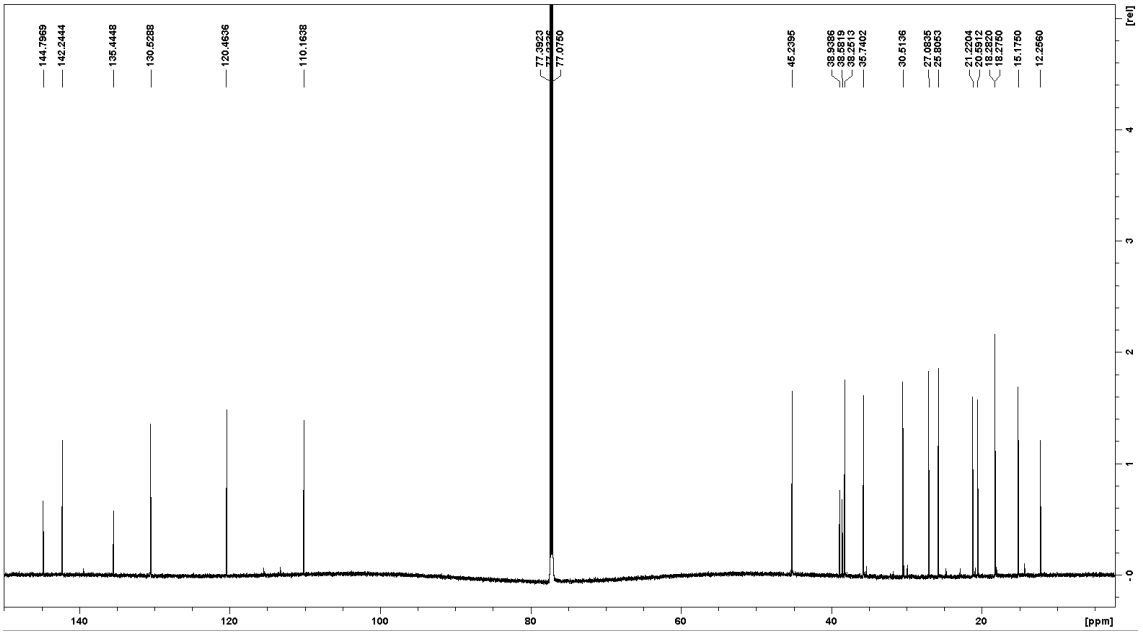


Fig. S37. (B) ^{13}C Spectrum of (5*S*,8*R*,9*R*,10*S*)-cleroda-3,12*E*,14-triene (**53**)

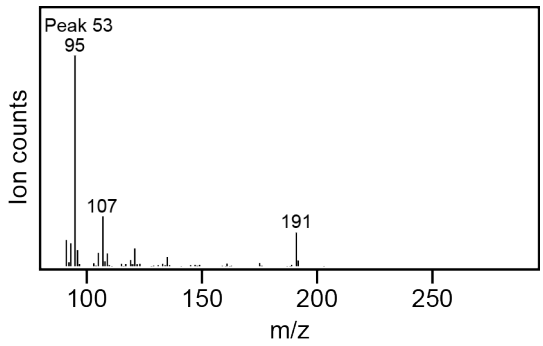
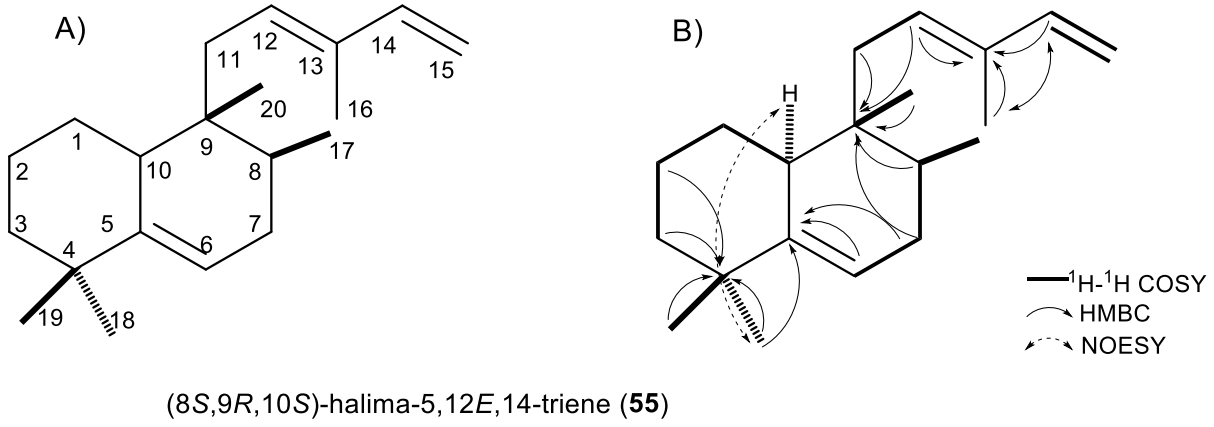


Fig. S38. MS of (5*S*,8*R*,9*R*,10*S*)-cleroda-3,12*E*,14-triene (**53**)



(8*S*,9*R*,10*S*)-halima-5,12*E*,14-triene (**55**)

Fig. S39. The product of ScLS upon co-expression with a DTC producing **16** (MtHPS). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-^{1}H COSY correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure.

Table S13. ^{1}H and ^{13}C NMR assignments for compound **55**, (8*S*,9*R*,10*S*)-halima-5,12*E*,14-triene (solvent CDCl3)

Position	(8*S*,9*R*,10*S*)-halima-5,12*E*,14-triene (**55**)	
	δ_H	δ_C
1a	1.77 (1H, m)	28.4
b	1.00 (1H, m)	
2a	1.56 (1H, m)	22.6
b	1.47 (1H, m)	
3a	1.37 (1H, m)	41.3
b	1.18 (1H, td, J = 13.2 Hz, 4.2 Hz)	
4		36.4

5		146.5
6	5.41 (1H, d, *J* = 4.4 Hz)	116.2
7a	1.80 (1H, m)	31.7
b	1.74 (1H, m)	
8	1.49 (1H, m)	34.1
9		38.6
10	2.06 (1H, d, *J* = 12.4 Hz)	40.7
11a	2.27 (1H, m)	35.5
b	2.02 (1H, m)	
12	5.53 (1H, t, *J* = 7.6 Hz)	129.8
13		135.5
14	6.38 (1H, dd, *J* = 17.3 Hz, 10.7 Hz)	142.2
15a	5.06 (1H, d, *J* = 17.3 Hz)	110.4
b	4.90 (1H, d, *J* = 10.7 Hz)	
16	1.75 (3H, s)	12.1
17	0.81 (3H, d, J = 6.8 Hz)	15.4
18	0.95 (3H, s)	29.0
19	1.03 (3H, s)	30.0
20	0.65 (3H, s)	16.1

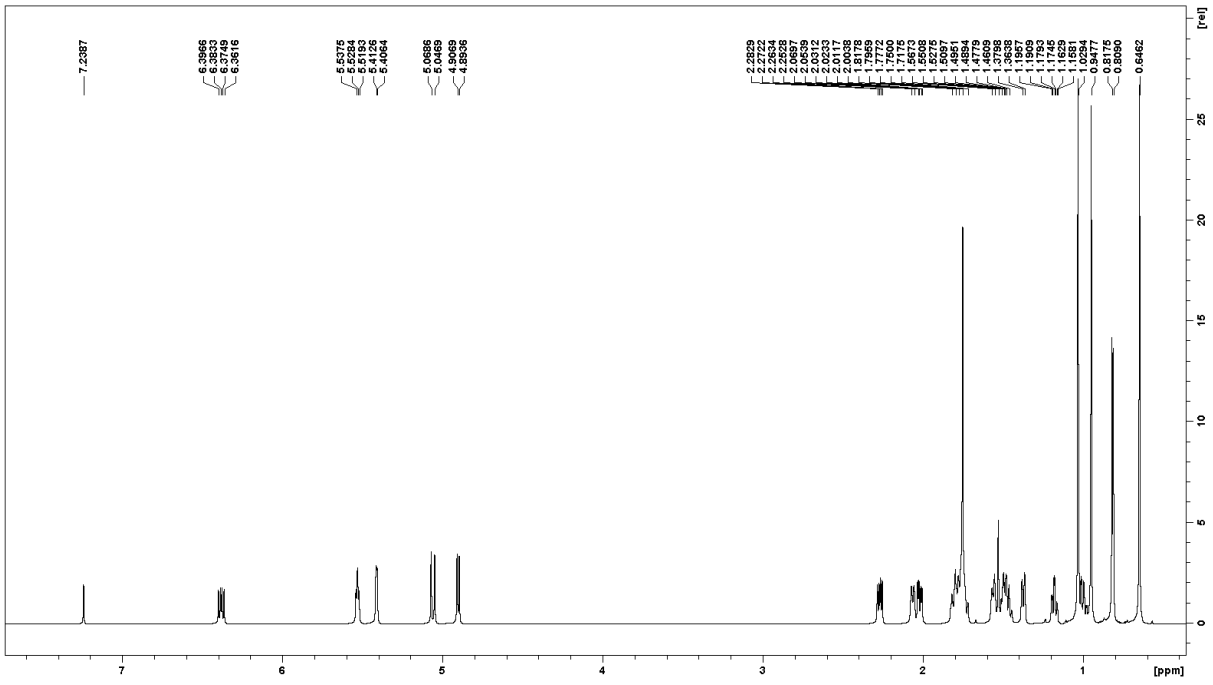


Fig. S40. (A) ^{1}H Spectrum of (8*S*,9*R*,10*S*)-halima-5,12*E*,14-triene (**55**)

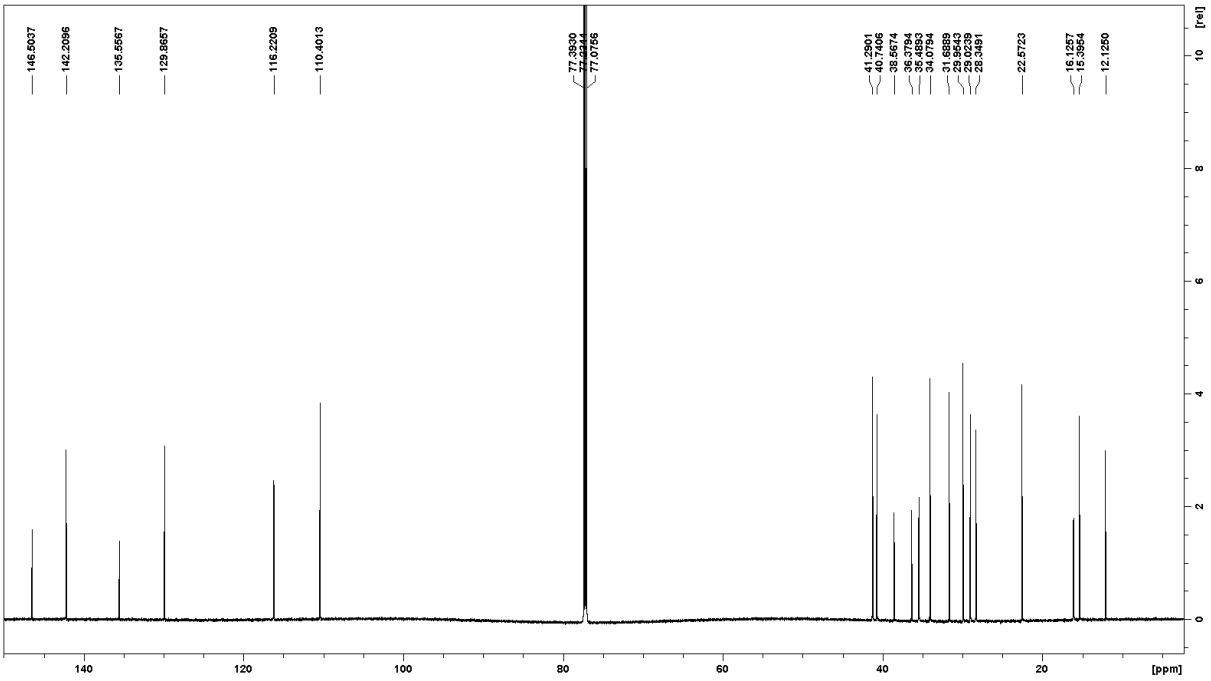


Fig. S40. (B) ^{13}C Spectrum of (8*S*,9*R*,10*S*)-halima-5,12*E*,14-triene (**55**)

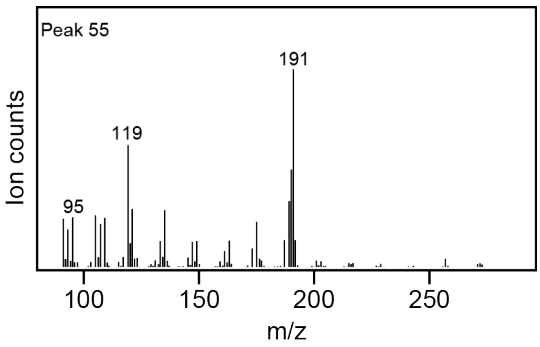
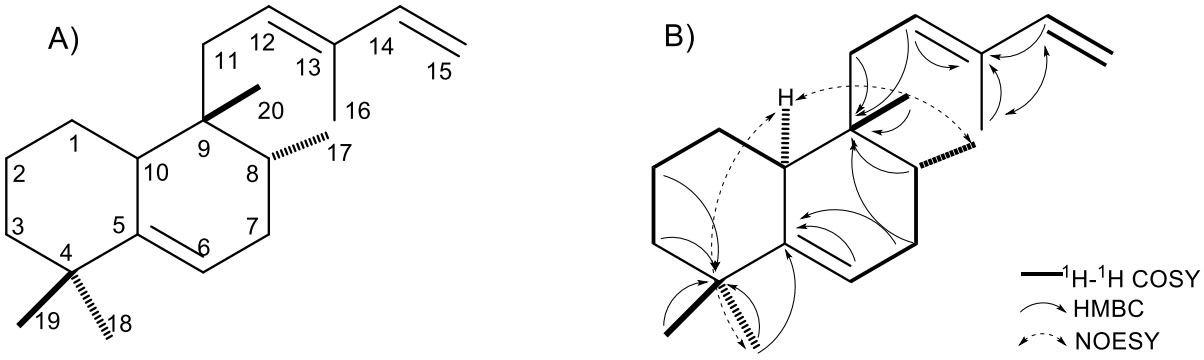


Fig. S41. MS of (8*S*,9*R*,10*S*)-halima-5,12*E*,14-triene (**55**)



(8*R*,9*R*,10*S*)-halima-5,12*E*,14-triene (**57**)

Fig. S42. The product of ScLS upon co-expression with the DTC producing **17** (OsCPS4:H561D). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-

^{1}H COSY correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure.

Table S14. ^{1}H and ^{13}C NMR assignments for compound **57**, (8*R*,9*R*,10*S*)-halima-5,12*E*,14-triene (solvent CDCl3)

Position	(8*R*,9*R*,10*S*)-halima-5,12*E*,14-triene (**57**)	
	δ_H	δ_C
1a	1.73 (1H, m)	29.5
b	1.03 (1H, m)	
2	1.58 (2H, m)	23.3
3a	1.41 (1H, d, J = 10.4 Hz)	42.7
b	1.15 (1H, td, J = 12.8 Hz, 5.3 Hz)	
4		37.0
5		147.2
6	5.34 (1H, brs)	114.9
7a	2.03 (1H, m)	31.6
b	1.73 (1H, m)	
8	1.61 (1H, m)	33.4
9		37.9
10	1.96 (1H, m)	41.7
11a	2.21 (1H, m)	32.4
b	1.98 (1H, m)	
12	5.54 (1H, t, *J* = 7.8 Hz)	130.9
13		135.3
14	6.39 (1H, dd, *J* = 17.3 Hz, 10.6 Hz)	142.4
15a	5.03 (1H, d, *J* = 17.3 Hz)	110.1
b	4.89 (1H, d, *J* = 10.6 Hz)	
16	1.69 (3H, s)	12.3
17	0.80 (3H, d, J = 6.7 Hz)	15.1
18	0.97 (3H, s)	26.8
19	1.02 (3H, s)	29.7
20	0.83 (3H, s)	22.8

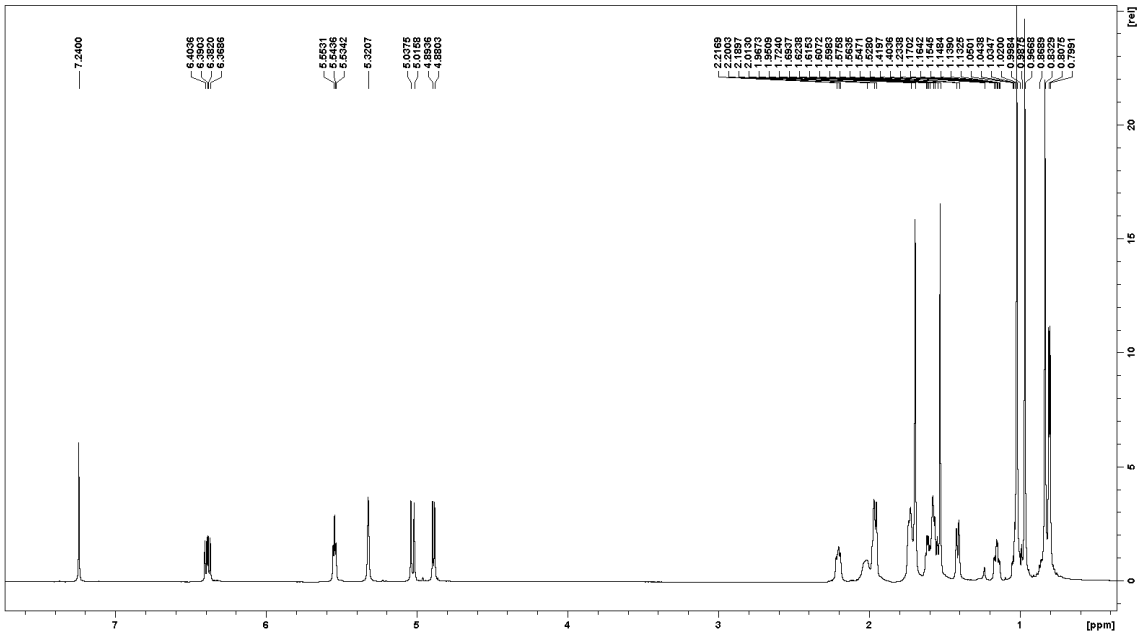


Fig. S43. (A) ^{1}H Spectrum of (8*R*,9*R*,10*S*)-halima-5,12*E*,14-triene (**57**)

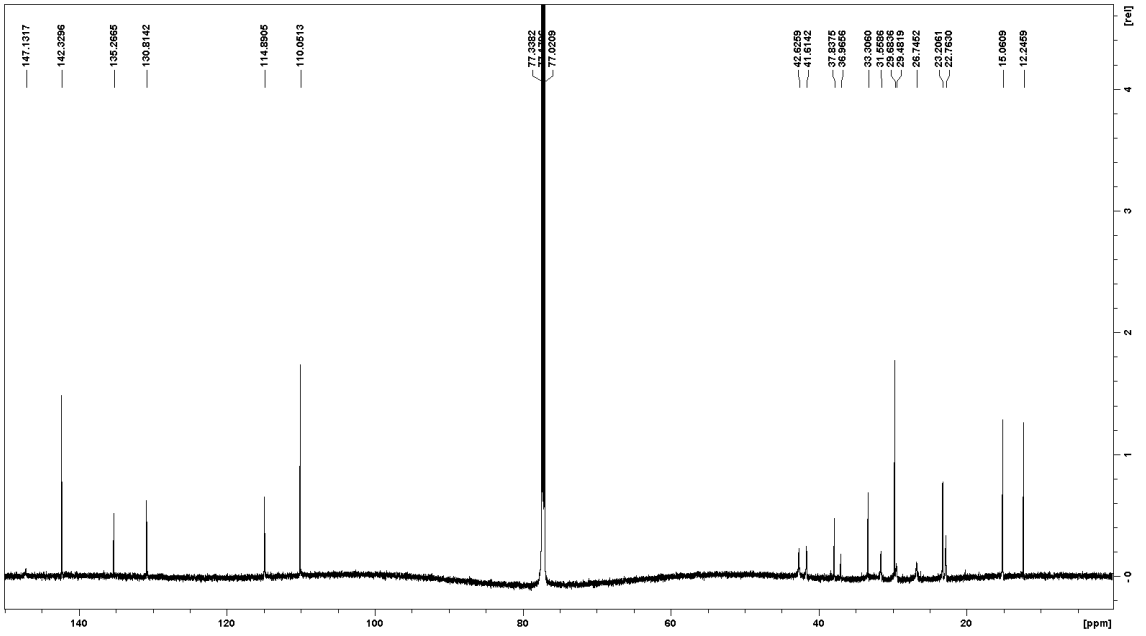


Fig. S43. (B) ^{13}C Spectrum of (8*R*,9*R*,10*S*)-halima-5,12*E*,14-triene (**57**)

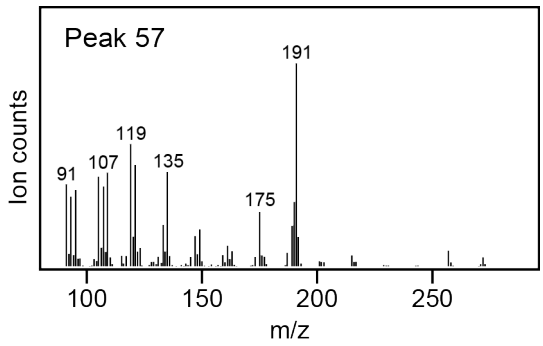
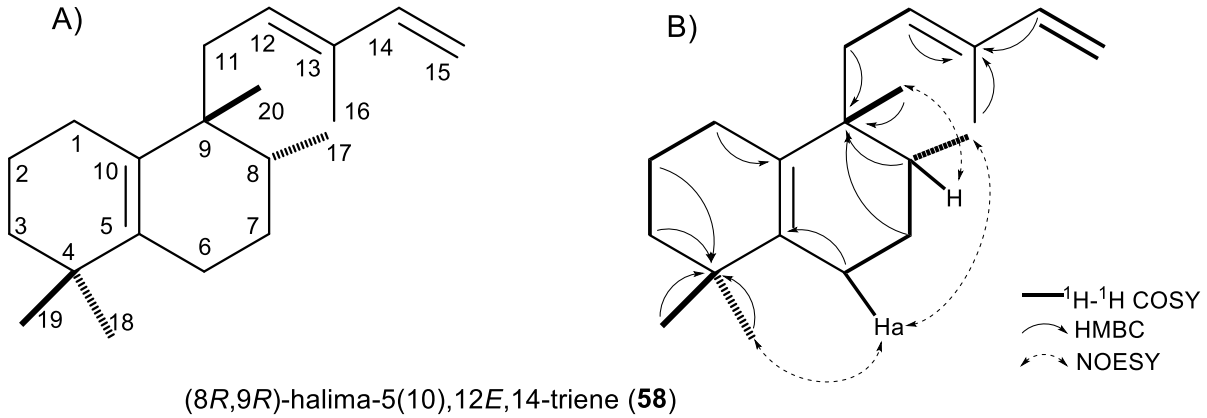


Fig. S44. MS of (8*R*,9*R*,10*S*)-halima-5,12*E*,14-triene (**57**)



(8*R*,9*R*)-halima-5(10),12*E*,14-triene (**58**)

Fig. S45. The product of ScLS upon co-expression with a DTC producing **18** (MvCPS1:W323F/F505Y). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-^{1}H COSY correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure.

Table S15. ^{1}H and ^{13}C NMR assignments for compound **58**, (8*R*,9*R*)-halima-5(10),12*E*,14-triene (solvent CDCl3)

Position	(8*R*,9*R*)-halima-5(10),12*E*,14-triene (**58**)	
	δ_H	δ_C
1	1.92 (2H, m)	26.5
2	1.54 (2H, m)	20.2
3	1.40 (2H, m)	40.1
4		34.6
5		135.7
6a	2.00 (1H, m)	23.8
b	1.89 (1H, m)	
7a	1.50 (1H, m)	27.5

b	1.38 (1H, m)	
8	1.47 (1H, m)	38.2
9		41.1
10		132.4
11a	2.25 (1H, dd, *J* = 15.7 Hz, 8.9 Hz)	34.1
b	2.05 (1H, dd, *J* = 15.7 Hz, 5.9 Hz)	
12	5.50 (1H, t, *J* = 7.2 Hz)	132.7
13		133.1
14	6.32 (1H, dd, *J* = 17.4 Hz, 10.7 Hz)	142.3
15a	5.02 (1H, d, *J* = 17.4 Hz)	110.0
b	4.87 (1H, d, *J* = 10.7 Hz)	
16	1.71 (3H, s)	12.2
17	0.86 (3H, d, *J* = 6.8 Hz)	16.6
18	0.99 (3H, s)	28.6
19	0.96 (3H, s)	28.6
20	0.99 (3H, s)	26.2

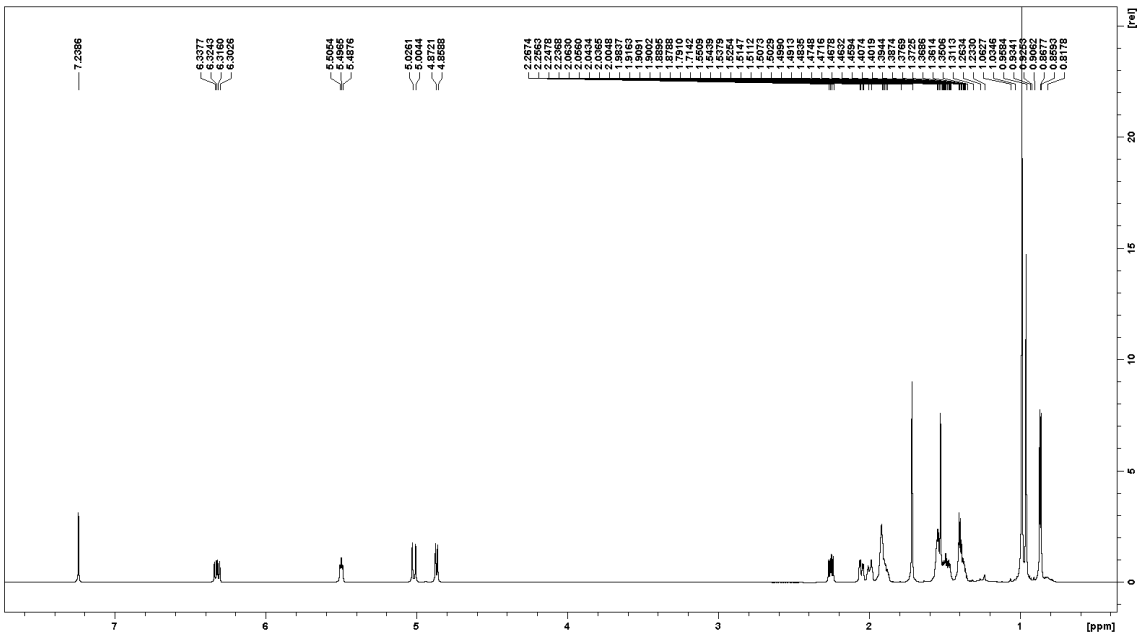


Fig. S46. (A) ^{1}H Spectrum of (8*R*,9*R*)-halima-5(10),12*E*,14-triene (**58**)

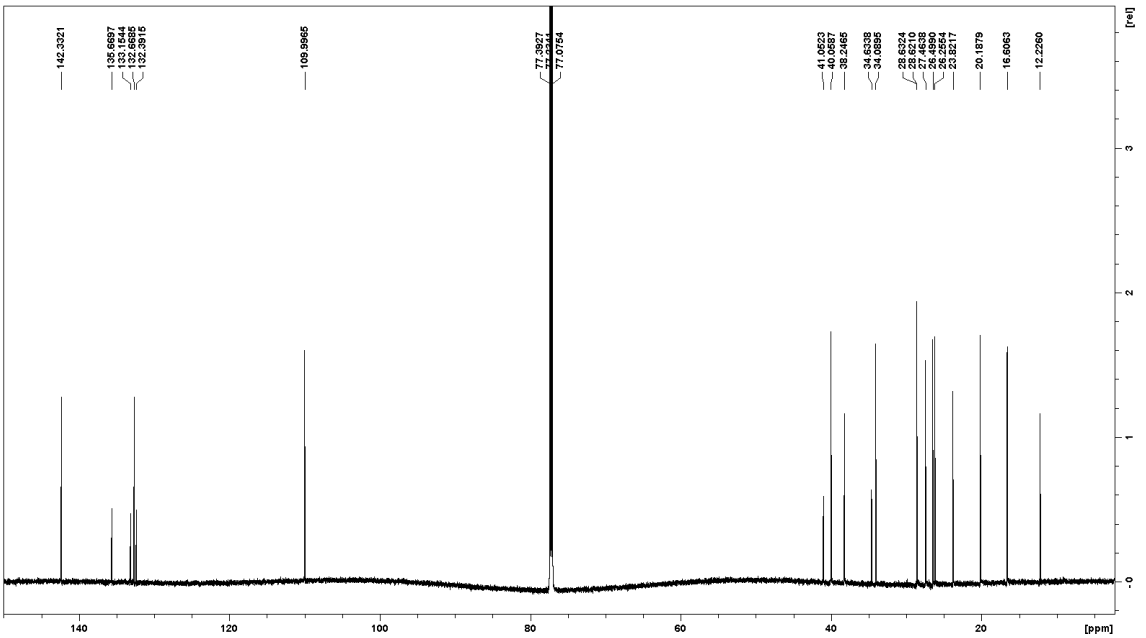


Fig. S46. (B) ^{13}C Spectrum of (8*R*,9*R*)-halima-5(10),12*E*,14-triene (**58**)

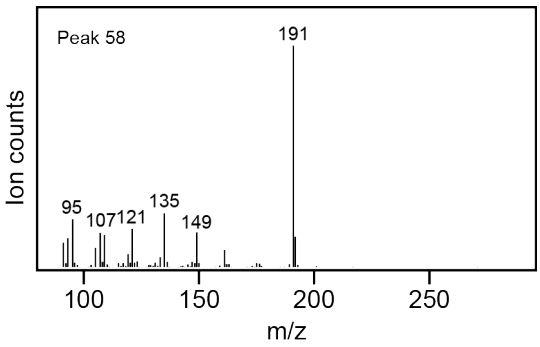


Fig. S47. MS of (8*R*,9*R*)-halima-5(10),12*E*,14-triene (**58**)

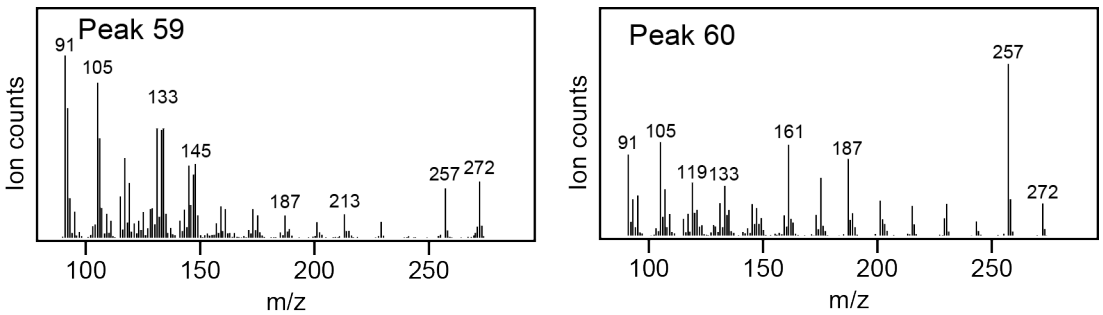


Fig. S48. Product identification by comparing both RT and MS from GC-MS analyses to authentic standards. For general reference the MS are presented here.

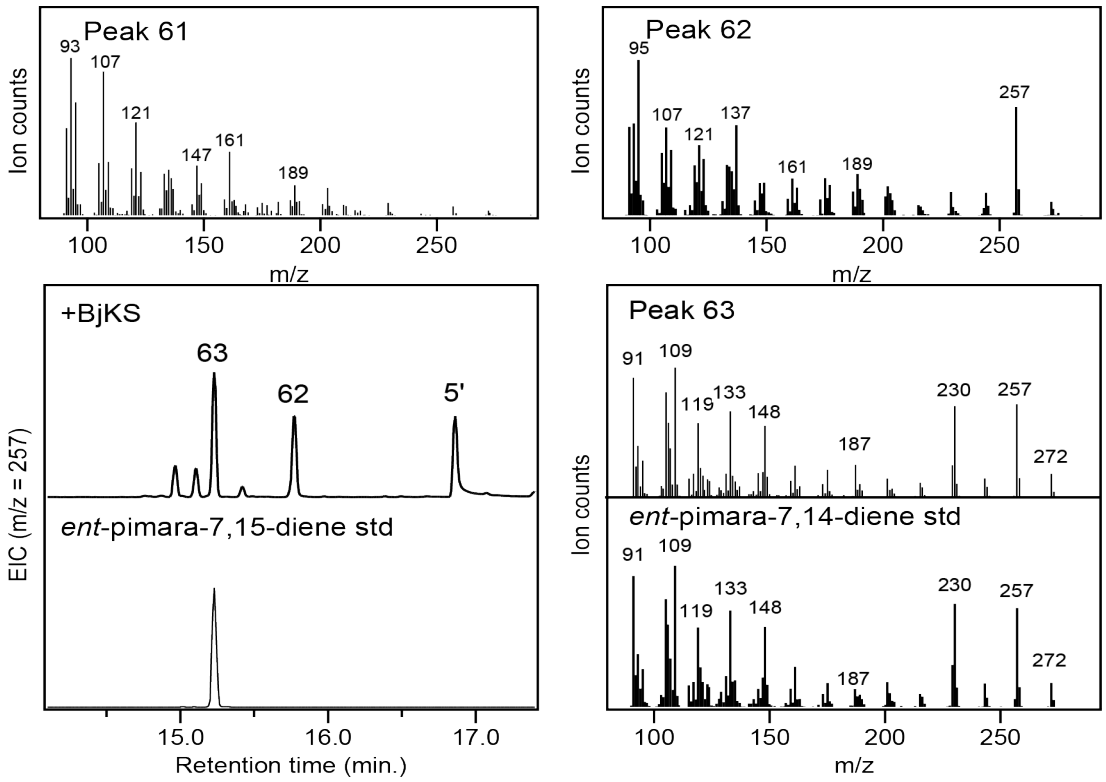


Fig. S49. Product identification by comparing both RT and MS from GC-MS analyses to authentic standards. New product **63** afforded by BjKS reacting with CPP (**5**) was identified by such comparison to the known *enantiomer*. Otherwise for general reference only MS are shown here.

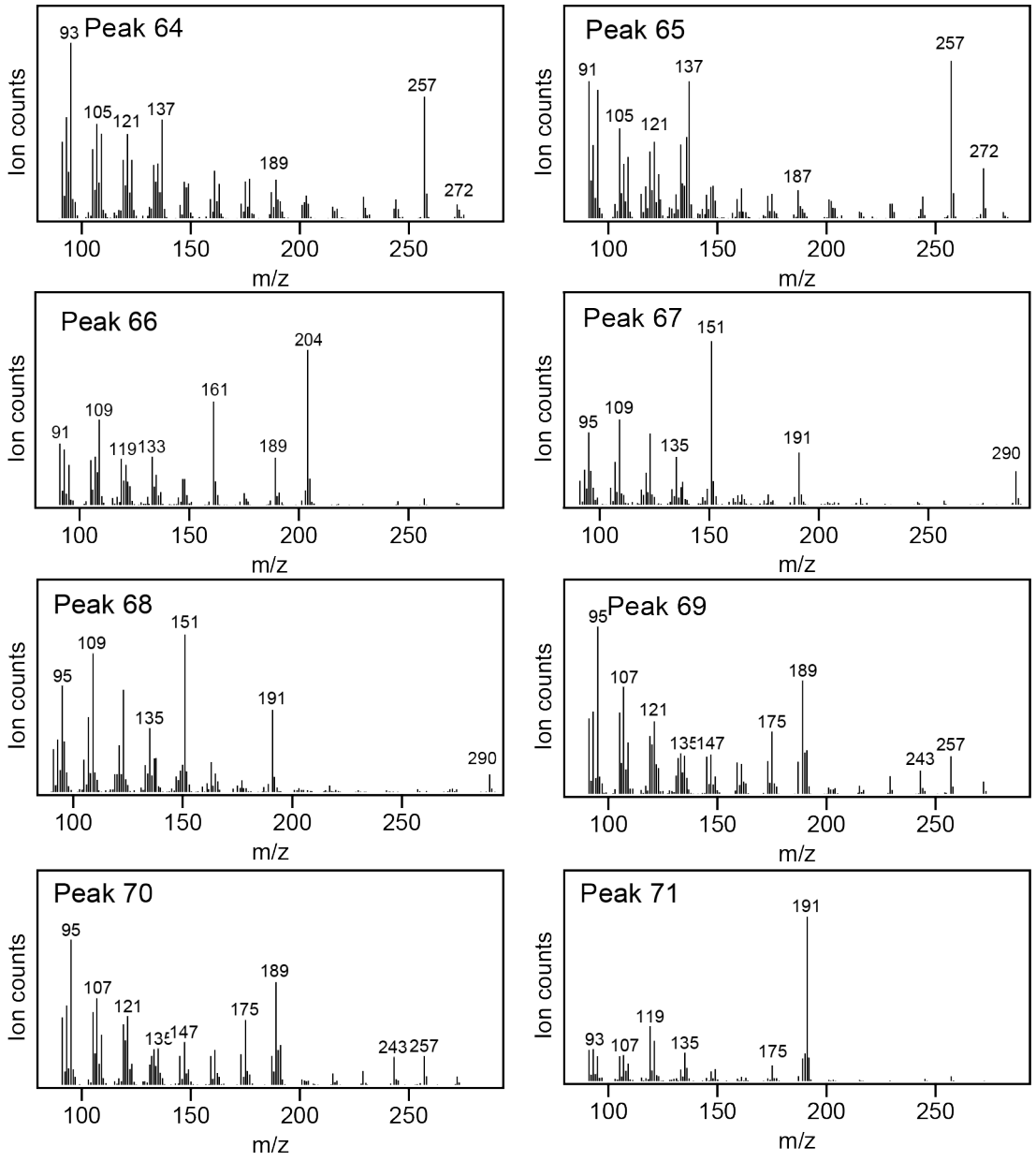


Fig. S50. Product identification by comparing both RT and MS from GC-MS analyses to authentic standards. For general reference the MS are presented here.

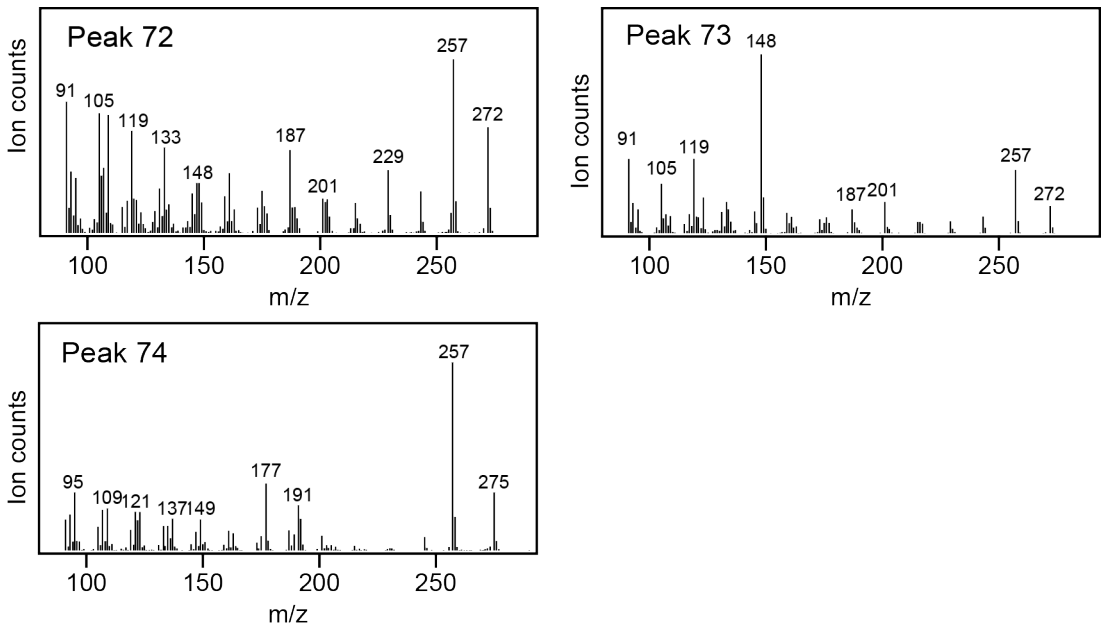
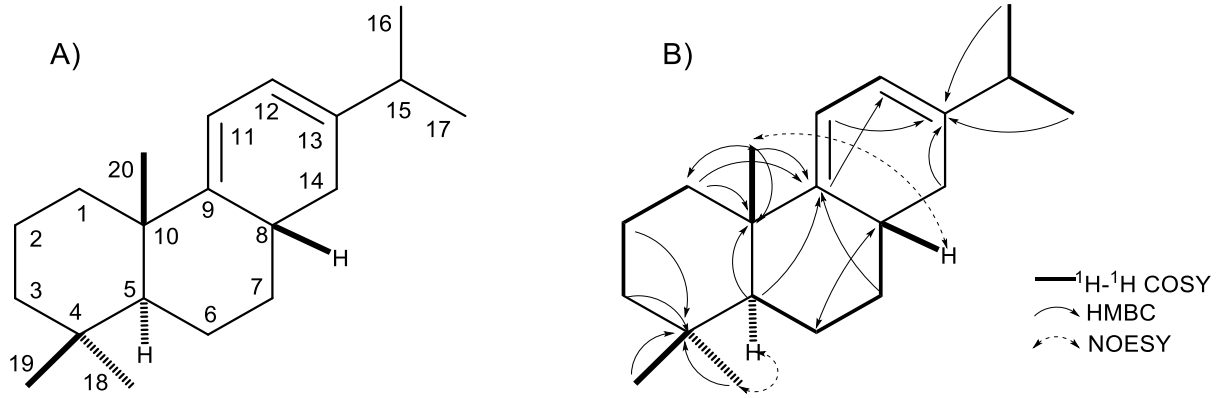


Fig. S51. Product identification by comparing both RT and MS from GC-MS analyses to authentic standards. For general reference the MS are presented here.



(8S,10S)-abieta-9(11),12-diene (**75**)

Fig. S52. The product of SmMS upon co-expression with the DTC producing **7** (OsCPS4). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-^{1}H COSY correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure.

Table S16. ^{1}H and ^{13}C NMR assignments for compound **75**, (8*S*,10*S*)-abieta-9(11),12-diene (solvent CDCl3).

Position	(8*S*,10*S*)-abieta-9(11),12-diene (**75**)	
	δ_H	δ_C
1a	1.79 (1H, d, *J* = 12.7 Hz)	37.2
b	1.41 (1H, m)	
2a	1.69 (1H, m)	19.0
b	1.57 (1H, m)	
3a	1.45 (1H, m)	42.1
b	1.20 (1H, td, *J* = 13.5 Hz, 4.4 Hz)	
4		33.6
5	1.07 (1H, m)	53.4
6a	1.70 (1H, m)	21.8
b	1.48 (1H, m)	
7a	2.03 (1H, m)	37.0
b	1.27 (1H, m)	
8	2.55 (1H, m)	33.0
9		150.5
10		39.4
11	5.64 (1H, dd, *J* = 6.0 Hz, 1.8 Hz)	112.6
12	5.61 (1H, d, *J* = 5.8 Hz)	115.9
13		141.9
14a	6.33 (1H, t, *J* = 15.4 Hz)	34.4
b	2.31 (1H, m)	
15	2.28 (1H, m)	34.2
16	1.06 (1H, d, *J* = 3.1 Hz)	21.2
17	1.05 (1H, d, *J* = 2.9 Hz)	21.0
18	0.91 (3H, s)	33.5
19	0.90 (3H, s)	21.9
20	1.08 (3H, s)	21.7

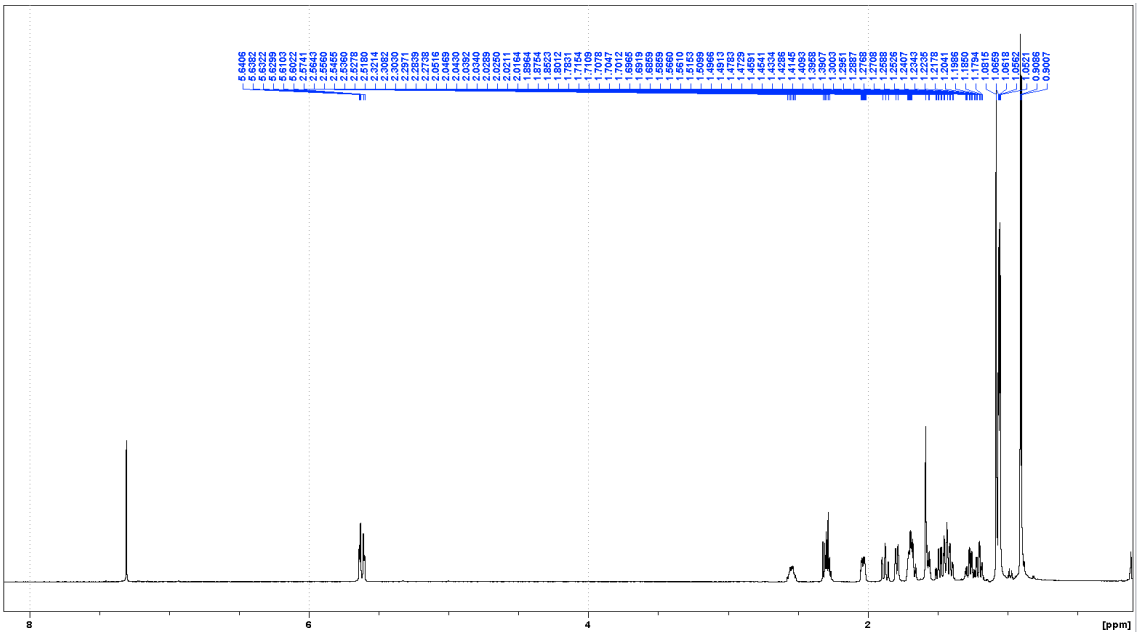


Fig. S53. (A) ^{1}H Spectrum of (8*S*,10*S*)-abieta-9(11),12-diene (**75**)

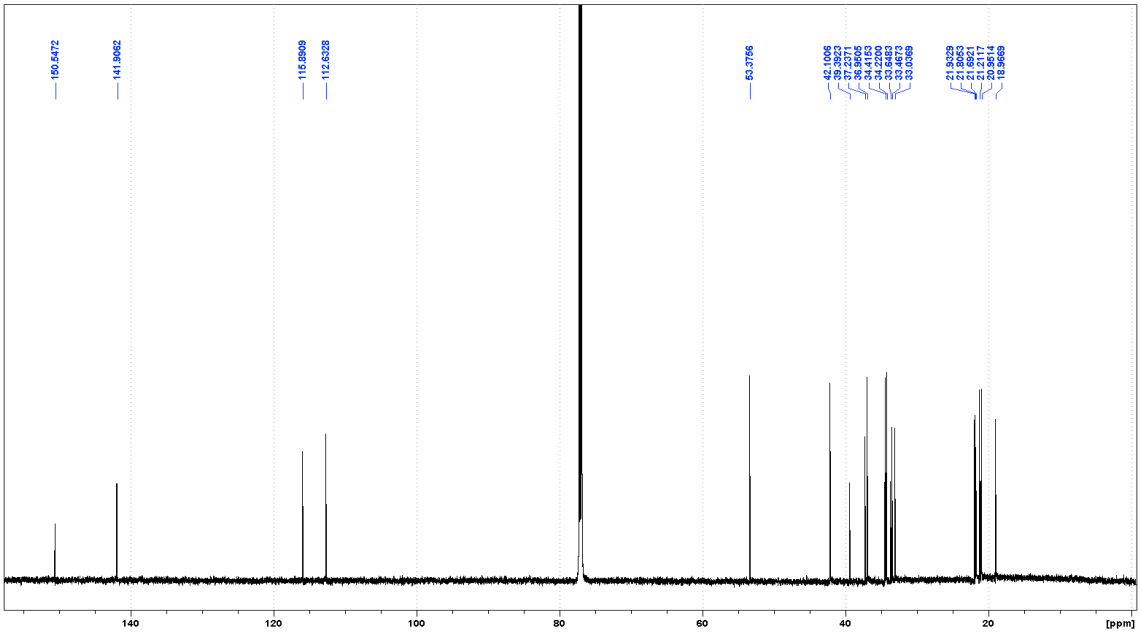


Fig. S53. (B) ^{13}C Spectrum of (8*S*,10*S*)-abieta-9(11),12-diene (**75**)

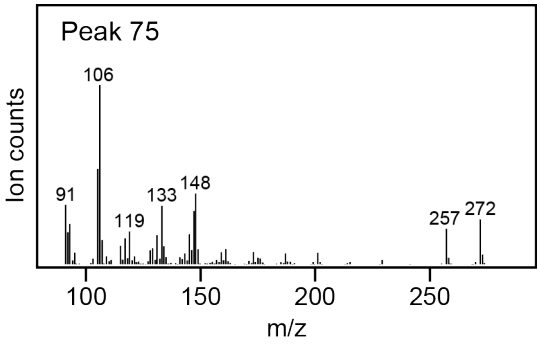
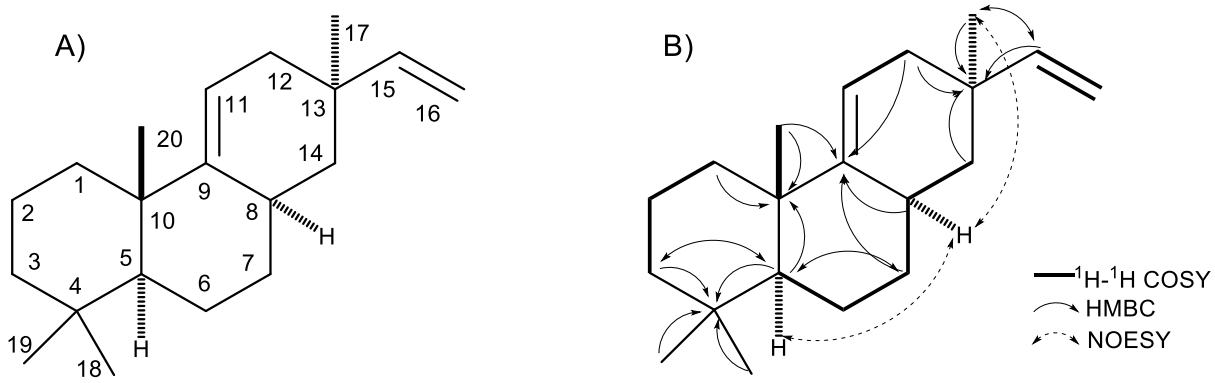


Fig. S54. MS of (8*S*,10*S*)-abieta-9(11),12-diene (**75**)



syn-pimara-9(11),15-diene (**76**)

Fig. S55. The product of SaPS upon co-expression with the DTC producing **7** (OsCPS4). Structure with (A) carbon numbering or (B) arrows indicating ^{1}H-^{1}H COSY correlations and selected HMBC and NOESY Nuclear Overhauser Effect dipole-dipole correlations used to assign the structure.

Table S17. ^{1}H and ^{13}C NMR assignments for compound **76**, *syn*-pimara-9(11),15-diene (solvent CDCl3).

Position	*syn*-pimara-9(11),15-diene (**76**)	
	δ_H	δ_C
1a	1.63 (1H, m)	37.9
b	1.32 (1H, m)	
2a	1.63 (1H, m)	19.3
b	1.52 (1H, m)	
3a	1.36 (1H, m)	42.5
b	1.14 (1H, m)	
4		33.9
5	0.94 (1H, m)	53.9
6a	1.62 (1H, m)	22.9
b	1.43 (1H, m)	
7a	1.87 (1H, m)	36.4
b	0.98 (1H, m)	
8	2.25 (1H, m)	31.4
9		150.0
10		39.5
11	5.26 (1H, brd, J = 6.0 Hz)	112.6
12a	2.03 (1H, brd, J = 16.8 Hz)	37.5

b	1.69 (1H, m)	
13		35.0
14a	1.46 (1H, m)	42.8
b	1.09 (1H, m)	
15	5.80 (1H, dd, *J* = 17.6 Hz, 10.8 Hz)	150.6
16a	4.91 (1H, d, *J* = 17.6 Hz)	109.4
b	4.85 (1H, d, *J* = 10.8 Hz)	
17	0.90 (3H, s)	22.2
18	0.84 (3H, s)	22.2
19	0.83 (3H, s)	33.6
20	1.01 (3H, s)	21.2

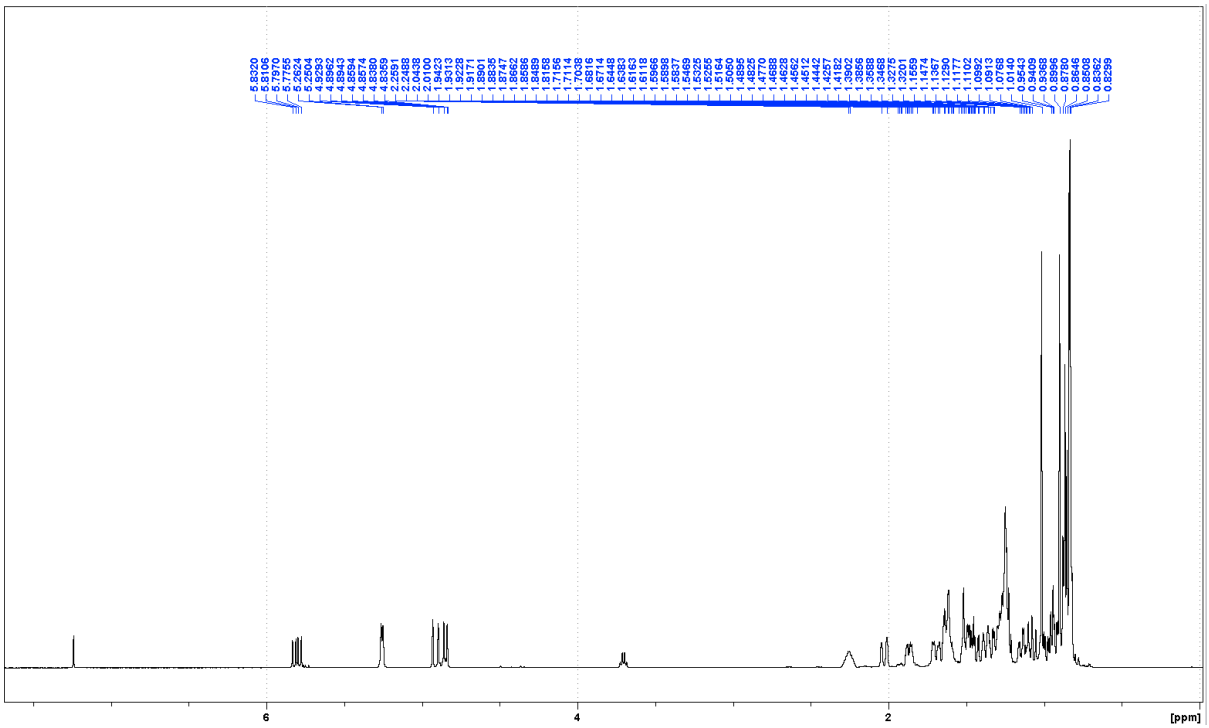


Fig. S56. (A) ^{1}H Spectrum of *syn*-pimara-9(11),15-diene (**76**)

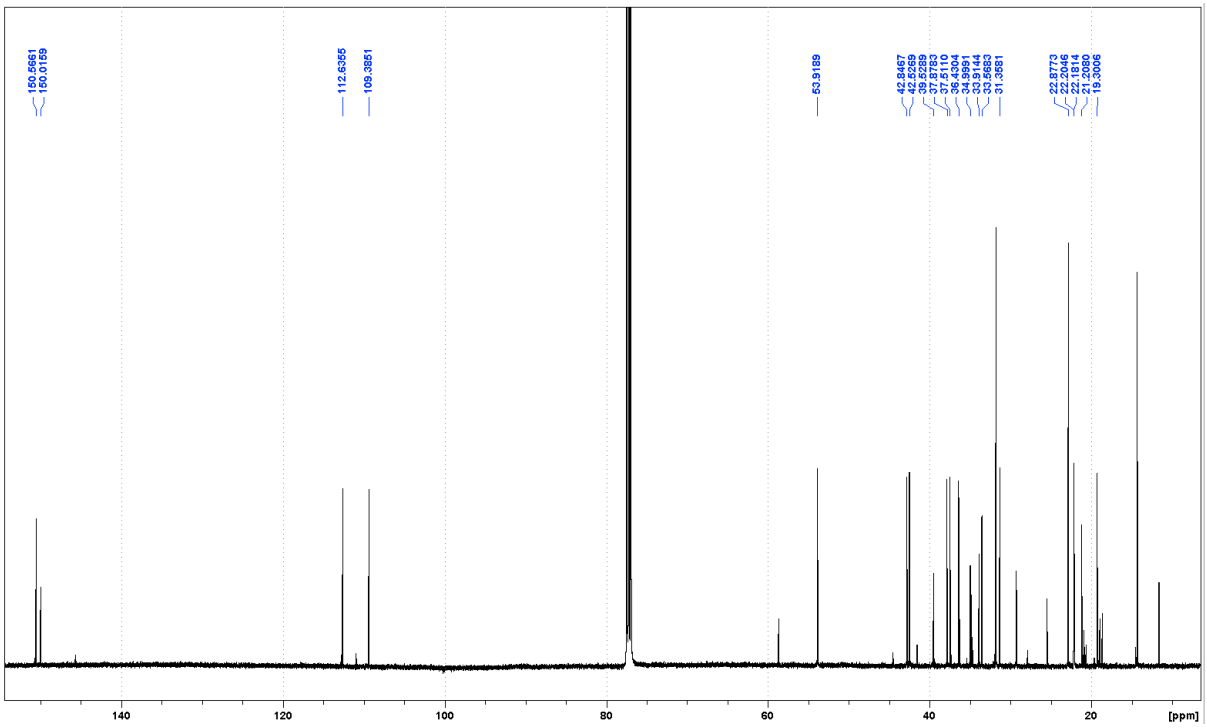


Fig. S56. (B) ^{13}C Spectrum of *syn*-pimara-9(11),15-diene (**76**)

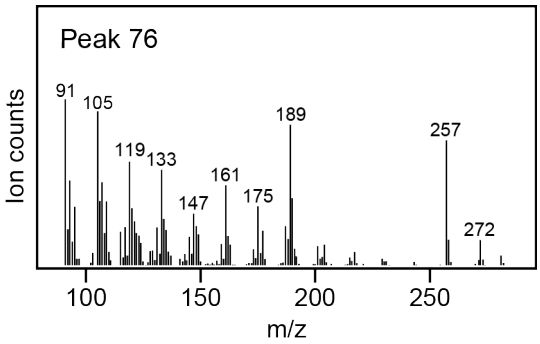


Fig. S57. MS of *syn*-pimara-9(11),15-diene (**76**)

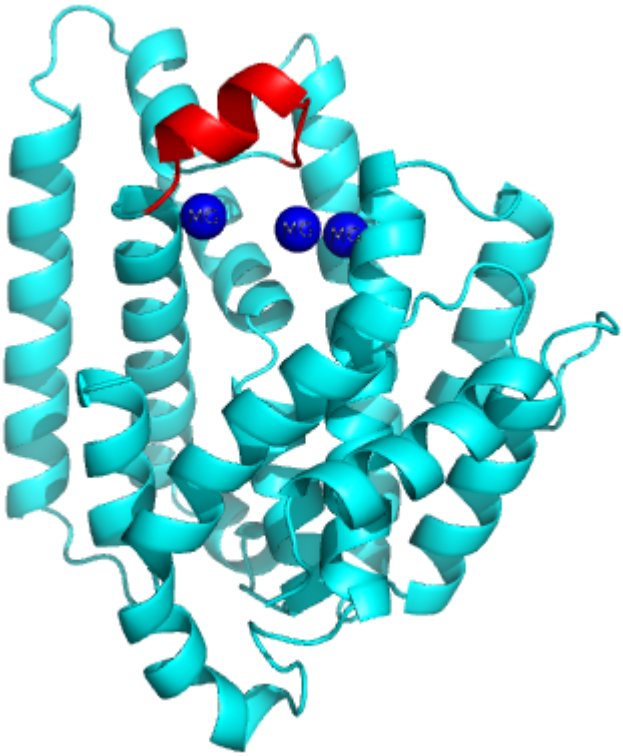


Fig. S58. Modeled BjKS structure for the 'closed' conformation obtained by adding the missing residues (211-220, highlighted in red) and the trinuclear Mg^{2+} ions (shown as blue spheres).

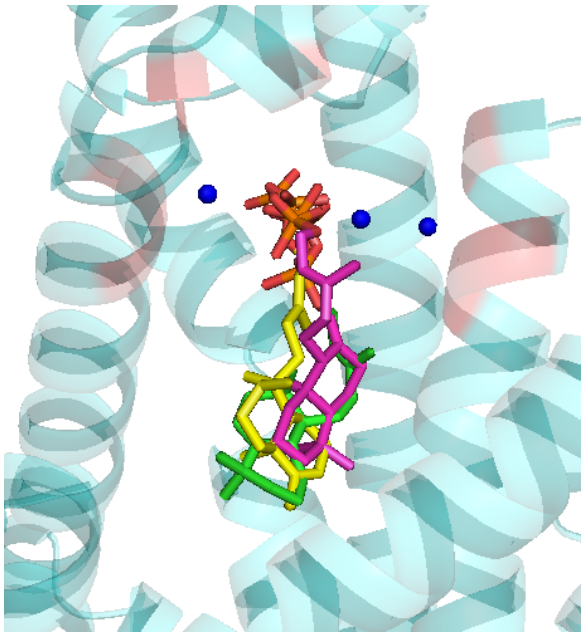


Fig. S59. Overlay of the three representative substrates after each was individually docked into the modeled 'closed' BjKS structure. All substrates are shown in stick with the labdane-type substrate *ent*-CPP (**6**) in purple, the clerodane-type substrate *syn*-KPP (**15**) in yellow and the halimane-type substrate *syn*-HPP (**17**) in green.

Table S18: Summary of the products obtained from this study with ratios shown in parenthesis when multiple products were observed.

Product[a]	Substrate[a]	DTS	Identification[b]
20	**3**	KgTS (26%) ScLS (14%)	(Nakano, *et al.*, 2010)
21	**3**	KgTS (31%) ScLS (22%)	(Nakano, *et al.*, 2010)
22	**3**	KgTS (43%) ScLS (41%)	(Nakano, *et al.*, 2010)
23	**3**	ScLS (23%) BjKS EtKS	commercial standard
24	**4**	KgTS (8%)	(Huang, *et al.*, 2017, Sato, *et al.*, 2013)
25	**4**	KgTS (84%)	This study
26	**4**	KgTS (8%)	This study
27	**2**	KgTS (73%) ScLS	This study
28	**2**	KgTS (27%)	(Jia, *et al.*, 2016)
28	**1**	SaPS BjKS (36%) EtKS (53%)	(Jia, *et al.*, 2016)
29	**9**	KgTS ScLS (69%) AtKS	This study
30	**9**	SsSS SaPS BjKS EtKS	This study
31	**18**	KgTS (44%) ScLS (12%)	This study
32	**18**	KgTS (56%) SsSS ScLS (13%) SaPS BjKS EtKS	This study
33	**19**	KgTS ScLS	This study
34	**10**	AbCAS	(Zerbe, *et al.*, 2012)
35	**5**	ScLS	(Yamada, *et al.*, 2016, Ikeda, *et al.*, 2016)
36	**8**	ScLS	(Yamada, *et al.*, 2016, Ikeda, *et*

			al., 2016)
37	**1**	ScLS	(Jia*, et al.*, 2016)
38	**5**	AbCAS (26%)	(Jia*, et al.*, 2016)
39	**5**	AbCAS (74%)	This study
40	**6**	ScLS (62%)	This study
41	**6**	ScLS (38%) AtKS OsKS BjKS EtKS	(Morrone*, et al.*, 2009)
42	**7**	ScLS (32%)	(Jia*, et al.*, 2016)
43	**7**	ScLS (68%)	(Jia*, et al.*, 2016)
44	**9**	ScLS (31%)	This study
45	**11**	ScLS SaPS (60%) AtKS OsKS BjKS EtKS	(Mafu*, et al.*, 2015)
46	**13**	ScLS	(Jia*, et al.*, 2016)
47	**10**	ScLS (39%) SmMS (78%) SaPS BjKS (46%) EtKS (84%)	(Mafu*, et al.*, 2015)
48	**10**	ScLS (36%) SmMS (22%) BjKS (54%) EtKS (16%)	(Mafu*, et al.*, 2015)
49	**10**	ScLS (25%)	This study
50	**12**	ScLS (67%)	(Heskes*, et al.*, 2018)
51	**12**	ScLS (33%)	(Heskes*, et al.*, 2018)
52	**14**	ScLS	This study
53	**15**	ScLS (83%)	This study
54	**15**	ScLS (17%)	(Jia*, et al.*, 2016)
55	**16**	ScLS	This study
56	**17**	ScLS (73%)	(Jia*, et al.*, 2016)
57	**17**	ScLS (27%)	This study
58	**18**	ScLS (75%)	This study
59	**5**	SmMS	(Gao*, et al.*, 2009)
60	**5**	SaPS	(Xu*, et al.*, 2014)
61	**1**	BjKS (64%) EtKS (47%)	(Jia*, et al.*, 2016)

62	**5**	BjKS (42%) EtKS (73%)	(Jia*, et al.*, 2016)
63	**5**	BjKS (58%) EtKS (27%)	This study
64	**6**	SaPS (65%)	(Jia*, et al.*, 2016)
65	**6**	SaPS (35%)	(Jia*, et al.*, 2016)
66	**8**	SaPS BjKS EtKS	(Jia*, et al.*, 2016)
67	**12**	SaPS BjKS (18%) EtKS (51%)	(Jia*, et al.*, 2016)
68	**12**	BjKS (82%) EtKS (49%)	(Jia*, et al.*, 2016)
69	**14**	BjKS (89%) EtKS (29%)	(Jia*, et al.*, 2016)
70	**14**	BjKS (11%) EtKS (71%)	(Jia*, et al.*, 2016)
71	**17**	SaPS BjKS EtKS	(Jia*, et al.*, 2016)
72	**7**	BjKS (80%) EtKS (75%)	(Wilderman*, et al.*, 2004)
73	**7**	SaPS (30%) BjKS (20%) EtKS (25%)	(Morrone*, et al.*, 2006)
74	**11**	SaPS (40%)	(Mafu*, et al.*, 2015)
75	**7**	SmMS	This study
76	**7**	SaPS (70%)	This study

[a]Products and substrates are numbered as defined in the text.

[b]Products were identified based on either 'previous' reports for these enzymes or GC-MS based 'comparison' to other previously reported DTS products (with accompanying reference), or were determined in 'this study' by NMR based structural analysis (or comparison to the characterized enantiomer).

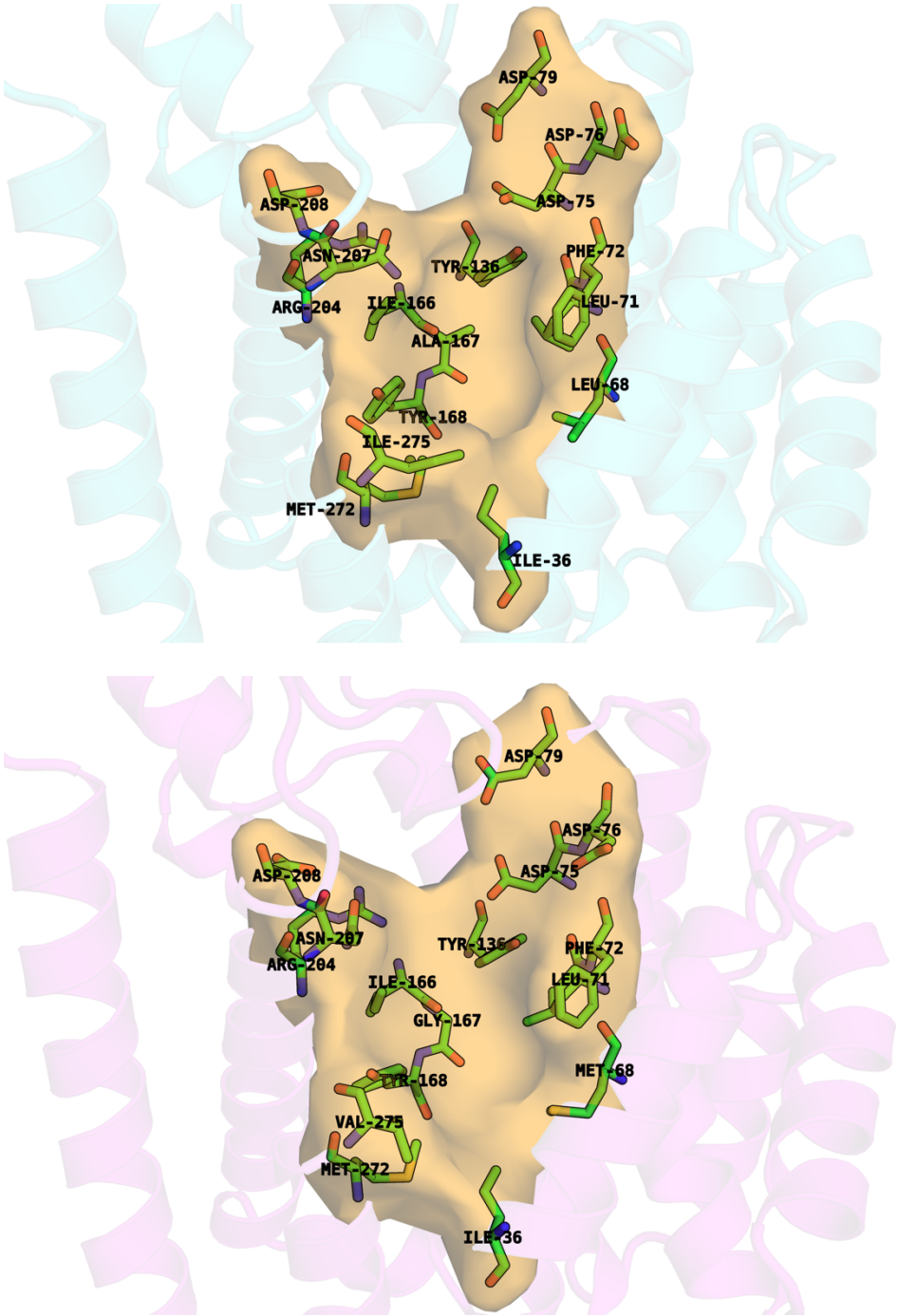


Fig. S60. Active sites for modeled BjKS (top, blue) and derived EtKS (bottom, magenta) structures (i.e., for the 'closed' conformation), as defined by indicated key residues.

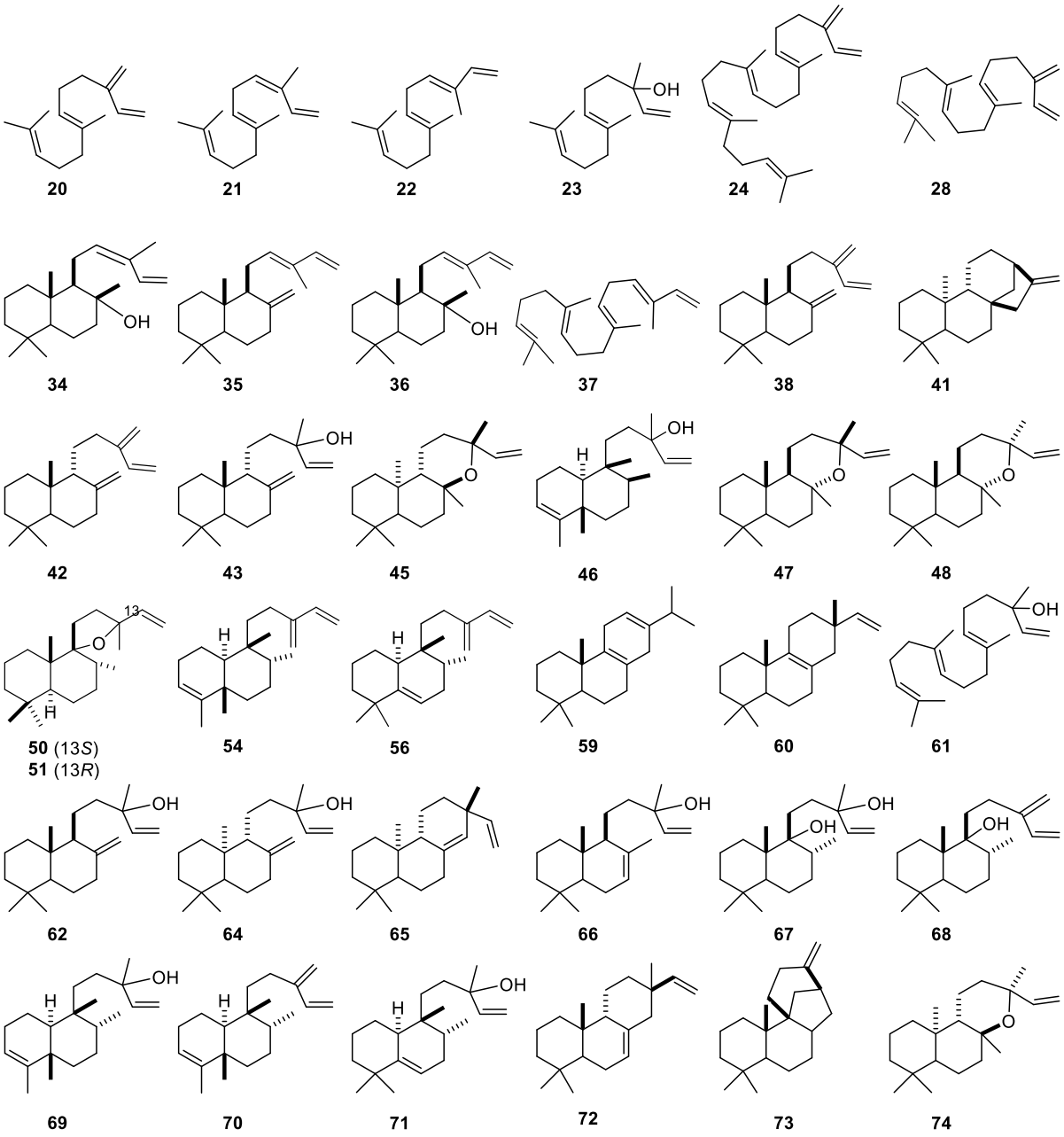


Fig. S61. Chemical structures of all the previously known DTSs products also observed here (numbering as defined in the text).

References

Johnson SR, Bhat WW, Bibik J, Turmo A, Hamberger B, Genomics Consortium EM, Hamberger B (2018) A database-driven approach identifies additional diterpene synthase activities in the mint family (Lamiaceae). *The Journal of biological chemistry*.

Pelot KA, Hagelthorn DM, Hong YJ, Tantillo DJ, Zerbe P (2019) Diterpene Synthase-Catalyzed Biosynthesis of Distinct Clerodane Stereoisomers. *Chembiochem : a European journal of chemical biology* 20(1):111-117.

Pelot KA, Chen R, Hagelthorn DM, Young CA, Addison JB, Muchlinski A, Tholl D, Zerbe P (2018) Functional Diversity of Diterpene Synthases in the Biofuel Crop Switchgrass. *Plant physiology* 178(1):54-71.

Nakano C, Hoshino T, Sato T, Toyomasu T, Dairi T, Sassa T (2010) Substrate specificity of the CYC2 enzyme from Kitasatospora griseola: production of sclarene, biformene, and novel bicyclic diterpenes by the enzymatic reactions of labdane- and halimane-type diterpene diphosphates. *Tetrahedron Letters* 51(1):125-128.

Huang AC, Kautsar SA, Hong YJ, Medema MH, Bond AD, Tantillo DJ, Osbourn A (2017) Unearthing a sesterterpene biosynthetic repertoire in the Brassicaceae through genome mining reveals convergent evolution. *Proceedings of the National Academy of Sciences of the United States of America* 114(29):E6005-E6014.

Sato T, Yamaga H, Kashima S, Murata Y, Shinada T, Nakano C, Hoshino T (2013) Identification of novel sesterterpene/triterpene synthase from Bacillus clausii. *Chembiochem : a European journal of chemical biology* 14(7):822-825.

Jia M, Potter KC, Peters RJ (2016) Extreme promiscuity of a bacterial and a plant diterpene synthase enables combinatorial biosynthesis. *Metabolic engineering* 37:24-34.

Zerbe P, Chiang A, Yuen M, Hamberger B, Hamberger B, Draper JA, Britton R, Bohlmann J (2012) Bifunctional cis-abienol synthase from Abies balsamea discovered by transcriptome sequencing and its implications for diterpenoid fragrance production. *The Journal of biological chemistry* 287(15):12121-12131.

Yamada Y, Komatsu M, Ikeda H (2016) Chemical diversity of labdane-type bicyclic diterpene biosynthesis in Actinomycetales microorganisms. *J Antibiot (Tokyo)* 69(7):515-523.

Ikeda H, Shin-Ya K, Nagamitsu T, Tomoda H (2016) Biosynthesis of mercapturic acid derivative of the labdane-type diterpene, cyslabdan that potentiates imipenem activity against methicillin-resistant Staphylococcus aureus: cyslabdan is generated by mycothiol-mediated xenobiotic detoxification. *J Ind Microbiol Biotechnol* 43(2-3):325-342.

Morrone D, Chambers J, Lowry L, Kim G, Anterola A, Bender K, Peters RJ (2009) Gibberellin biosynthesis in bacteria: separate ent-copalyl diphosphate and ent-kaurene synthases in Bradyrhizobium japonicum. *FEBS letters* 583(2):475-480.

Mafu S, Potter KC, Hillwig ML, Schulte S, Criswell J, Peters RJ (2015) Efficient heterocyclisation by (di)terpene synthases. *Chemical communications*.

Heskes AM, Sundram TCM, Boughton BA, Jensen NB, Hansen NL, Crocoll C, Cozzi F, Rasmussen S, Hamberger B, Hamberger B, Staerk D, Moller BL, Pateraki I (2018) Biosynthesis of bioactive diterpenoids in the medicinal plant Vitex agnus-castus. *The Plant journal : for cell and molecular biology* 93(5):943-958.

Gao W, Hillwig ML, Huang L, Cui G, Wang X, Kong J, Yang B, Peters RJ (2009) A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. *Org Lett* 11(22):5170-5173.

Xu M, Hillwig ML, Lane AL, Tiernan MS, Moore BS, Peters RJ (2014) Characterization of an orphan diterpenoid biosynthetic operon from Salinispora arenicola. *Journal of natural products* 77(9):2144-2147.

Jia M,Peters RJ (2016) Extending a single residue switch for abbreviating catalysis in plant ent-kaurene synthases. *Frontiers in Plant Science* In press.

Wilderman PR, Xu M, Jin Y, Coates RM, Peters RJ (2004) Identification of syn-pimara-7,15-diene synthase reveals functional clustering of terpene synthases involved in rice phytoalexin/allelochemical biosynthesis. *Plant physiology* 135(4):2098-2105.

Morrone D, Jin Y, Xu M, Choi SY, Coates RM, Peters RJ (2006) An unexpected diterpene cyclase from rice: functional identification of a stemodene synthase. *Archives of biochemistry and biophysics* 448(1-2):133-140.